Desh

Volume 40 April 1954 Number 4

OF MICHIGAN
MAY 10 1954
PERIODICAL
READING ROOM

Proceedings of the

National Academy of Sciences

of the United States of America

The Proceedings of the National Academy of Sciences

OFFICERS OF THE ACADEMY

DETLEV W. BRONK
President

GEORGE W. CORNER
Vice President

ROGER ADAMS
Foreign Secretary

ALEXANDER WETMORE
Home Secretary

WILLIAM J. ROBBINS
Treasurer

EDITORIAL BOARD OF THE PROCEEDINGS

LINUS PAULING
Chairman

ALEXANDER WETMORE, *Home Secretary*
EDWIN B. WILSON, *Managing Editor*

ROGER ADAMS, *Foreign Secretary*
WILLIAM W. RUBEY, *Chairman of the National Research Council*

J. W. BEAMS
R. E. CLELAND
W. W. RUBEY
HARLOW SHAPLEY
F. E. TERMAN

G. W. BEADLE
E. A. DOISY
G. C. EVANS
C. H. GRAHAM
J. T. PATTERSON

GREGORY BREIT
L. C. DUNN
W. F. LIBBY
J. VON NEUMANN
PAUL WEISS

Subscription price is $7.50 per volume (12 issues); the price of a single issue is $.75. Make all remittances payable to The University of Chicago Press in U.S. currency or its equivalent. Authorized agent for the British Commonwealth, except North America and Australasia: Cambridge University Press, Bentley House, 200 Euston Road, London, N. W. 1, England.

Editorial correspondence should be addressed to the PROCEEDINGS of the NATIONAL ACADEMY of SCIENCES, Attention Miss Mary D. Alexander, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois.

Business correspondence should be addressed to the Journals Division, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois.

Subscribers are requested to notify The University of Chicago Press and their local postmaster immediately of change of address.

Microfilms of complete volumes of this journal are available to regular subscribers only and may be obtained at the end of the volume year from University Microfilms, 313 N. First Street, Ann Arbor, Michigan.

Entered as second-class matter June 21, 1920 at the Post-office at Easton, Pennsylvania, under the Act of August 24, 1912. Accepted for mailing at a special rate of postage provided for in Section 1103, Act of October 3, 1917. Authorized June 18, 1920.

PRINTED IN THE U. S. A.

THE PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
is published for
THE NATIONAL ACADEMY OF SCIENCES
by
THE UNIVERSITY OF CHICAGO PRESS

Proceedings of the

NATIONAL ACADEMY OF SCIENCES

Volume 40 • Number 4 • April 15, 1954

*THE INHIBITION OF HYDROGENASE BY NITRIC OXIDE**

By Alvin I. Krasna† and D. Rittenberg‡

DEPARTMENT OF BIOCHEMISTRY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY

Communicated February 26, 1954

The enzyme hydrogenase is of particular interest because of the comparative simplicity of its substrate. In its activation of molecular hydrogen, hydrogenase simulates the action of active platinum catalysts.[1] Although hydrogenase has not been obtained in pure form, considerable information is available concerning its properties. The evidence indicates that hydrogenase contains ferrous iron;[2, 3] that it forms with carbon monoxide an inactive complex,[3, 4] which is photochemically dissociated by light;[3, 5] and that it can be oxidized to a ferric analogue, which can form a complex with cyanide.[3, 4] Hydrogenase reacts with hydrogen according to the following equation, where E represents the enzyme:[6]

$$E + H_2 \rightleftarrows H{:}E^- + H^+. \qquad (1)$$

Fisher, Krasna, and Rittenberg[7] have found that hydrogenase can be reversibly oxygenated and deoxygenated:

$$E + O_2 \rightleftarrows EO_2. \qquad (2)$$

In support of the view that hydrogenase is an iron-containing enzyme we now present evidence that it combines with and is inhibited by nitric oxide. Iron compounds, both inorganic and organic, are known to form complexes with nitric oxide.[8]

Experimental.—Proteus vulgaris was grown, and cell-free extracts were prepared as described in a previous paper.[6] The enzymatic activity was determined by measuring the rate of the exchange reaction:[3, 6]

$$H_2 + HDO \rightleftarrows HD + H_2O.$$

The activity is expressed as the first-order rate constant, k.

Tank nitric oxide was purified according to Farkas and Melville.[9] Oxygen-free hydrogen was prepared by passing tank hydrogen over hot copper. Nitric oxide-hydrogen mixtures containing from 0.1 to 10 per cent nitric oxide were prepared volumetrically. This procedure was not suitable for the preparation of very dilute mixtures, since traces of oxygen in the system would convert some or all of the nitric oxide to nitrogen dioxide. To prepare mixtures containing less than 0.1 per cent nitric oxide, a mixture of 10 per cent NO and 90 per cent H_2 was cooled to the temperature of liquid or solid nitrogen for 1 hour. At these temperatures all the nitrogen dioxide is frozen out, and the vapor pressure of solid nitric oxide determines the final pressure of nitric oxide. These gas mixtures were kept in contact with the bacterial suspension for 10 minutes. The concentration of nitric oxide

in the mixtures of low concentration was determined by mass spectrometric analysis.

Results.—The effect of different concentrations of nitric oxide on the hydrogenase activity of suspensions of *Proteus vulgaris* is shown in Table 1. Nitric oxide is a powerful inhibitor of hydrogenase. At low nitric oxide concentrations the inhibition can be partly reversed by removal of the NO-H_2 mixture by evacuation and its replacement with H_2. At high concentrations (1 per cent NO and above), the inhibition could not be reversed even by the addition of sodium hydrosulfite, which reduces nitric oxide to nitrous oxide.[10] The suspensions which remained inactive after removal of the NO did not catalyze the ortho-para hydrogen conversion.[6] The active suspension carried out this conversion with a rate constant, k, of 4.6×10^{-3} per minute. Nitric oxide also inhibited the hydrogenase activity of a cell-free extract.

Nitric oxide is neither oxidized nor reduced by the bacterial suspension, since the composition of the NO-H_2 mixture over the bacteria does not change detectably in 24 hours.

TABLE 1*

INHIBITION OF HYDROGENASE ACTIVITY BY NITRIC OXIDE

NITRIC OXIDE CONCENTRATION		IN PRESENCE OF NO		AFTER REMOVAL OF NO†	
Volume Per Cent in Gas	Molarity in Solution‡	Rate Constant $k \times 10^3$	Per Cent Inhibition	Rate Constant $k \times 10^3$	Per Cent Inhibition
0	0	1.1	...	1.1	...
0.002...	4×10^{-6}	0.14	87	...	...
0.008...	1.6×10^{-7}	0.08	93	0.71	35
0.1.....	2×10^{-6}	0.03	97	0.66	40
1.0.....	2×10^{-5}	0.00	100	0.07	94
10.....	2×10^{-4}	0.00	100	0.00	100

* Each flask contained 0.1 ml. of a bacterial suspension (0.16 mg. N), in a total volume of 5 cc. of 10 per cent D_2O buffered to pH 6.7.

† The NO was removed by evacuation, and the flask was filled with hydrogen.

‡ These values are calculated on the assumption that the solubility of NO in the buffer system is the same as in water.

Discussion.—Hydrogenase appears to be sporadically distributed among the various species of microörganisms.[11] It is likely that every organism which produces or utilizes hydrogen contains hydrogenase. Incubation of *Scenedesmus* and certain other photosynthetic organisms under hydrogen results in the appearance of hydrogenase activity and concomitant photoreduction:[12, 13, 14]

$$12H_2 + 6CO_2 \rightarrow C_6H_{12}O_6 + 6H_2O.$$

Photoreduction is abolished by oxygen, and the change from normal photosynthesis to photoreduction is inhibited by cyanide and other agents which are known to form complexes with ferric iron. Since it is likely that photoreduction and photosynthesis are closely related processes, it seems reasonable that hydrogenase is derived from a part of the photosynthetic apparatus concerned with the splitting of water. There are but few biological substances which can combine reversibly with oxygen; all these contain iron or copper. Hydrogenase has been shown to combine reversibly with oxygen and to contain iron. This suggests that hydrogenase is closely related to the compounds which transport oxygen in the higher phyla.

We suggest that in the evolutionary process hydrogenase may have differentiated to yield the oxygen-producing system of the photosynthetic mechanism, the oxygen-transporting system, hemoglobin, and the enzymes, catalase, peroxidase, and the cytochromes.

There is considerable evidence that the early atmosphere of the earth was rich in hydrogen and hydrocarbons and contained only traces of oxygen and carbon dioxide.[15] In such an environment hydrogenase can be effective in the catalysis of reductive reactions which consume hydrogen and yield energy. Examples would be the reduction of CO_2 to CH_4 or of oxygen to water. Both reactions are exergonic.

In the evolution of life, photoreduction probably preceded photosynthesis. The development of the photosynthetic mechanism, with the concomitant appearance of oxygen, inhibited hydrogenase and replaced the reducing reactions by oxidizing reactions as a source of energy. Hydrogenase is conceivably an evolutionary remnant.

Summary.—Nitric oxide at low concentrations inhibits hydrogenase activity. The relation of hydrogenase to the other iron-containing enzymes and to hemoglobin is discussed.

* This work was aided by a contract between the Office of Naval Research, Department of the Navy, and Columbia University (ONR-26602). The deuterium oxide was obtained on allocation of the Atomic Energy Commission.

† National Science Foundation Predoctoral Fellow, 1953–54. This report is from a dissertation to be submitted by Alvin I. Krasna in partial fulfilment of the requirement for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University.

‡ On leave, 1953–54. Present address: Institute for Advanced Study, Princeton, N.J.

[1] D. E. Green and L. H. Stickland, *Biochem. J.*, **28**, 898–900 (1934).

[2] W. S. Waring and C. H. Werkman, *Arch. Biochem.*, **4**, 75–87 (1944).

[3] H. Hoberman and D. Rittenberg, *J. Biol. Chem.*, **147**, 211–227 (1943).

[4] L. A. Hyndman, R. H. Burris, and P. W. Wilson, *J. Bact.*, **65**, 522–530 (1953).

[5] W. Kempner and F. Kubowitz, *Biochem. Zeitschr.*, **265**, 245–252 (1933).

[6] A. I. Krasna and D. Rittenberg, *J. Amer. Chem. Soc.* (in press).

[7] H. H. Fisher, A. I. Krasna, and D. Rittenberg, *J. Biol. Chem.* (in press).

[8] N. V. Sidgwick, *The Chemical Elements and Their Compounds*, Vol. 2. London: Oxford University Press, 1952.

[9] A. Farkas and H. W. Melville, *Experimental Methods in Gas Reactions*. London: Macmillan & Co., Ltd., 1939.

[10] Nitrous oxide does not inhibit hydrogenase.

[11] L. Farkas and E. Fischer, *J. Biol. Chem.*, **167**, 787–805 (1947).

[12] H. Gaffron, *Nature*, **143**, 204–205 (1939).

[13] H. Gaffron, *J. Gen. Physiol.*, **26**, 195–217 (1942).

[14] H. Gaffron, *Biol. Rev.*, **19**, 1–20 (1944).

[15] H. C. Urey, *The Planets: Their Origin and Development*. New Haven: Yale University Press, 1952.

A BIOCHEMICAL AND PHARMACOLOGICAL SUGGESTION ABOUT CERTAIN MENTAL DISORDERS

By D. W. Woolley and E. Shaw

ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK

Communicated February 16, 1954

Recent findings in this laboratory[1, 2] and elsewhere[3, 4] have permitted an understanding of some aspects of mental diseases in relation to the hormone-like compound, serotonin. Furthermore, these findings lead directly to a suggestion for a logical treatment of diseases known as "schizophrenia." The experimental observations have been made solely on laboratory animals, but they have reached a point where clinical trials in human psychiatric patients are required to test the validity of the conclusions. Being only biochemists, we are unable to do these experiments on patients and can only hope that this paper will stimulate those who are professionally qualified to undertake in man what we cannot pursue further in laboratory animals.

This discussion is going to revolve about serotonin,[5] or enteramine, if Erspamer's terminology is followed.[6] Serotonin is one of the latest hormone-like substances to be discovered. Chemically it has been shown to have the structure given in Figure 1.[7] Being a simple molecule, it has been synthetically produced without great difficulty[8, 9] and is thus readily available. It was discovered because it is the vasoconstrictor long known to form in the serum when blood clots. Erspamer's work[6] with it, done independently of that of Rapport, *et al.*,[5] on the vasoconstrictor material, was based on the abundant occurrence of this new compound in the enterochromaffinic cells of the gastric and intestinal mucosa. Serotonin has now been isolated from several different organs, including the brain, of a wide variety of animal forms.[10–12] There can be no doubt of its wide-spread occurrence in living things. Furthermore, a variety of pharmacological properties in isolated organs and tissues has been demonstrated, in addition to the vasoconstrictor effect.[6, 13] The pharmacological properties which have been described thus far are attributable in large part to the ability of serotonin to cause various smooth muscles to contract; but there are indications that it has other effects which may become more clearly defined as time for study of them goes by.

Aside from the occurrence and importance of serotonin, a second general idea is prerequisite to this discussion. It is now well known that several classes of drugs are related chemically to individual hormones and other essential metabolites.[14, 15] In fact, a major part of the pharmacological effects of these drugs is attributed to a specific interference with the biological functioning of these metabolites to which the drugs are related structurally. That is, the drugs are antimetabolites.

Our thinking and experimentation about certain mental disorders has been in the following vein. Several synthetic compounds have been produced which are very closely related in structure to serotonin. These were shown to antagonize, in a competitive fashion, the contractions of artery walls caused by serotonin.[16, 17] Also, the fact that the ergot alkaloids are structurally related to serotonin was appreciated, and it was demonstrated that several of these actually did antagonize

the action of serotonin on artery walls.[1, 2] A derivative of the ergot alkaloids, namely, lysergic acid diethylamide (cf. Fig. 1), was then shown by Gaddum, *et al.*,[3] to act as an antimetabolite of serotonin on smooth muscles. It is well known that this amide calls forth in man mental disturbances resembling those of schizophrenia.[18] From these experiences we therefore conceived the idea that the mental disturbances caused by lysergic acid diethylamide were to be attributed to an interference with the action of serotonin in the brain. The failure of ergotamine or of ergotoxin to produce similar mental aberrations was postulated to be due to the failure of these more complex derivatives of lysergic acid to penetrate into the brain. Gaddum also was cognizant of the mental effects of lysergic acid diethylamide and of the occurrence of serotonin in the brain. We have surmised that he has been thinking, just as we have, about the relationship of serotonin to the mental disturbances induced by the drug.

Even before the study of the ergot alkaloids as antimetabolites of serotonin, another type of alkaloid, viz., yohimbine, was recognized to be structurally similar to serotonin and was actually demonstrated to function as an antimetabolite of it in the artery-wall test.[1, 2] Then two members of yet a third class of alkaloid—the harmala alkaloids—were seen to be structural analogues of serotonin and were actually shown to antagonize its action on artery walls.[2]

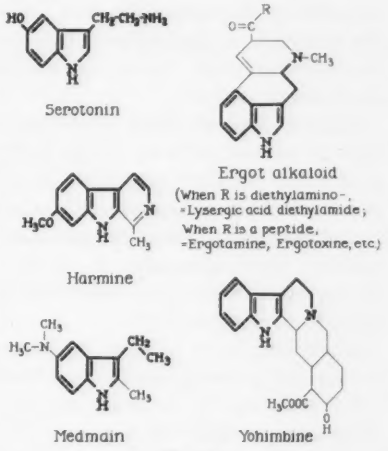

FIG. 1.

Although the chemical structures of the ergot alkaloids, yohimbine, and the harmala alkaloids are not very similar, all three classes of drugs have in common a resemblance to serotonin (Fig. 1). In each case the hormone has been modified in a different way in order to arrive at the particular class of alkaloid. Furthermore, they have all been shown to function as antagonists to the hormone.

The important point for the present discussion is that they all have something else in common. This is the ability of at least some members of each group to cause mental disturbances. If one does no more than read a textbook of pharmacology, one is struck by this fact.[19] Witness the production of schizophrenia with simple amides of lysergic acid, the hallucinations and euphoria called forth by some of the harmala alkaloids, and the mental disturbances attributed to the use of yohimbine. Perhaps even the aboriginal use of crude yohimbine as an aphrodisiac is further small evidence.

Bearing in mind the mental disturbances caused by alkaloids known to be antagonists of serotonin, let us examine other types of compounds which are not alkaloids but merely synthetic serotonin antagonists. What do we find? The highly active synthetic antimetabolite of serotonin, medmain (cf. Fig. 1),[20] when injected into mice, calls forth convulsions much resembling those of human epilepsy. Hydrazino-phthalazine, which is said by Taylor, Page, and Corcoran[4] to antagonize some

of the biological effects of serotonin, has been found to produce in man mental aberrations somewhat similar to those of schizophrenia.[21]

The thesis of this paper is that these pharmacological findings indicate that serotonin has an important role to play in mental processes and that the suppression of its action results in a mental disorder. In other words, it is the lack of serotonin which is the cause of the disorder. If now a deficiency of serotonin in the central nervous system were to result from metabolic rather than from pharmacologically induced disturbances, these same mental aberrations would be expected to become manifest. Perhaps such a deficiency is responsible for the natural occurrence of the diseases.

Of course, this may be the wrong inference from the pharmacological findings. It must be recognized that the classes of alkaloids and synthetic drugs we have been discussing are the very ones which, at one time or another, have been used for the reduction of high blood pressure. The mental disturbances may be no more than the reflection of diminished blood flow which they may cause in the brain. To this objection there is as yet no adequate answer, and it may be that it is the true explanation.

If the hypothesis about serotonin deficiency is accepted, then the obvious thing to do is to treat patients having appropriate mental disorders with serotonin. It is here that one must be a clinician rather than a biochemist; for how is one to induce, or even more to diagnose, schizophrenia in laboratory animals?

One may ask whether the administration of serotonin to humans is a safe procedure. More information is certainly needed on this point, but the few reports in the literature, where it has been done, as well as the toxicological studies in animals indicate no great risk.[22-24]

If serotonin is to be given to selected mental patients, it will probably have to be injected intramuscularly or subcutaneously. What assurance is there that it will find its way to the brain, even though it were to escape destruction by amine oxidase and possibly other enzymes in the periphery? The answer is that we have no assurance that it will reach the central nervous system or persist there. Most experience in experimental animals would, in fact, indicate that serotonin is rapidly destroyed, especially when given by the intravenous route. We have made estimations (by chemical assay) of the amount of serotonin found in the brain of a mouse 10 minutes after the intraperitoneal injection of 5 mg. of it, and have found less than 0.01 mg. Nevertheless, because mice are not necessarily the same as men in this respect, the clinical trial of serotonin in mental disorders may still be worth while.

If destruction or failure to penetrate the central nervous system should render abortive attempts to treat schizophrenia with intramuscular serotonin, then the idea should not be abandoned until the biochemists have made one more effort. This is so to alter the structure of the hormone that it will remain long-acting in the body and will penetrate into the brain from the peripheral circulation. This may not be an impossible task. There are already several indications that such desirable properties can be built into metabolites such as serotonin.

In summary, the suggestions we wish to make are the following: (1) serotonin probably plays a role in maintaining normal mental processes; (2) metabolically induced deficiency of serotonin may contribute to the production of some mental

disorders; (3) serotonin or a long-acting derivative of it may prove capable of alleviating disorders similar to schizophrenia.

[1] D. W. Woolley and E. Shaw, *Federation Proc.*, **12**, 293 (1953).
[2] E. Shaw and D. W. Woolley, *J. Biol. Chem.*, **203**, 979 (1953).
[3] J. H. Gaddum, C. O. Hebb, A. Silver and A. A. B. Swan, *Quart. J. Exper. Physiol.*, **38**, 255 (1953).
[4] R. D. Taylor, I. H. Page and A. C. Corcoran, *Arch. Int. Med.*, **88**, 1 (1951).
[5] M. M. Rapport, A. A. Green and I. H. Page, *J. Biol. Chem.*, **176**, 1243 (1948).
[6] V. Erspamer and B. Asero, *Nature*, **169**, 800 (1952).
[7] M. M. Rapport, *J. Biol. Chem.*, **180**, 961 (1949).
[8] K. E. Hamlin and F. E. Fischer, *J. Amer. Chem. Soc.*, **73**, 5007 (1951).
[9] M. E. Speeter, P. V. Heinzelman and D. I. Weisblat, *J. Amer. Chem. Soc.*, **73**, 5515 (1951).
[10] V. Erspamer, *Ricerca sc.*, **22**, 1568 (1952).
[11] V. Erspamer and M. Vialli, *Ricerca sc.*, **22**, 1420 (1952).
[12] B. M. Twarog and I. H. Page, *Amer. J. Physiol.*, **175**, 157 (1953).
[13] G. Reid and M. Rand, *Nature*, **169**, 801 (1952).
[14] A. J. Clark, *Handbuch der experimentellen Pharmakologie*, Vol. **4**. Berlin: Julius Springer, 1937.
[15] D. W. Woolley, *A Study of Antimetabolites*. New York: John Wiley & Sons, 1952.
[16] D. W. Woolley and E. Shaw, *J. Amer. Chem. Soc.*, **74**, 2948 (1952).
[17] D. W. Woolley and E. Shaw, *J. Biol. Chem.*, **203**, 69 (1953).
[18] W. A. Stoll, *Schweiz. Arch. Neurol. Psychiat.*, **68**, 279 (1947).
[19] T. Sollmann, *A Manual of Pharmacology*. Philadelphia: W. B. Saunders Co., 1949.
[20] E. Shaw and D. W. Woolley, *J. Pharm. and Exper. Therap.* (in press) (1954).
[21] M. Moser, J. Syner, S. Malitz and T. W. Mattingly, *J. Amer. Med. Assoc.*, **152**, 1329 (1953).
[22] I. H. Page and J. W. McCubbin, *Circ. Res.*, **1**, 354 (1953).
[23] G. Reid, *J. Physiol.*, **118**, 435 (1952).
[24] V. Erspamer, *Ricerca sc.*, **22**, 694 (1952).

DYNAMIC PROGRAMMING AND A NEW FORMALISM IN THE CALCULUS OF VARIATIONS

BY RICHARD BELLMAN

Communicated by John von Neumann, February 9, 1954

1. *Introduction.*—In a series of papers,[1-11] we have treated a number of mathematical problems arising from multistage decision processes. Problems of this type occur in the theory of probability;[1, 4, 5, 9] in mathematical economics;[1, 2, 6, 7, 8, 11] in control processes;[2, 8] in learning processes;[4, 9] and in many other fields as well.

In this paper we wish to show that the functional-equation technique introduced in the above works may be used to provide a new approach to some classical problems in the calculus of variations. In addition to furnishing a new analytic weapon, we feel that the method has great potentialities as a computational tool. As we have pointed out previously,[1, 3, 4] this approach seems ideally suited to the handling of variational problems involving stochastic processes. This point will be further enlarged upon in some forthcoming publications.

To illustrate the approach in its simplest setting, we shall consider first the problem of maximizing $\int_0^t F(x, z)\, du$, subject to the constraint $dx/du = G(x, z)$, $x(0) =$

c, where the analytic detail is at a minimum, and then turn to the eigenvalue problem derived from

$$\frac{d^2x}{dt^2} + \lambda\phi(t)x = 0, \qquad x(0) = x(1) = 0, \tag{1.1}$$

which is rather less straightforward.

In both cases we shall derive a partial differential equation for the desired quantity. The method we sketch below is also applicable to multidimensional eigenvalue problems whenever the underlying space has certain symmetry properties. Further results, together with a detailed exposition of the results contained herein, will appear in another publication.

2. *The Maximization Problem.*—Let us now consider the maximization problem described in the first section. We shall consistently assume that all functions that appear possess all the differentiability properties required to make our operations legitimate, since we are primarily interested here in presenting the basic formalism. The essence of our method lies in considering the maximum value as a function of the parameters which describe the state of the process, the "state variables." In this case these are c and t. We write

$$\operatorname*{Max}_{z} \int_0^t F(x, z)\, du = f(c, t). \tag{2.1}$$

The classical approach regards c and t as fixed and considers $\int_0^t F(x, z)\, du$ as a functional of z. In contrast, we regard c and t as parameters and seek to determine $z(0)$ as a function of c and t.

Speaking in terms of a multistage process, in place of determining the optimal continuation from one fixed position, we try to find the optimal first step from any position.

Let us proceed to the analysis which will clarify the above general remarks. We have

$$f(c, s + t) = \operatorname*{Max}_{z} \int_0^{s+t} F(x, z)\, du = \operatorname*{Max}_{z} \left[\int_0^s F(x, z)\, du + \int_s^{s+t} F(x, z)\, du\right]. \tag{2.2}$$

It is clear that whatever the choice of $z(u)$ in $0 \leq u \leq s$, if we wish to maximize the right side of (2.2) we must choose $z(u)$ in $s \leq u \leq s + t$ so as to maximize $\int_s^{s+t} F(x, z)\, du$ subject to the constraints $dx/dt = G(x, z)$, $x(s) = c(s)$. Here $c(s)$ is the new initial value obtained from the differential equation $dx/dt = G(x, z)$, $x(0) = c$.

In view of the invariant nature of the problem with respect to translations, we see that we must have

$$\int_s^{s+t} F(x, z)\, du = f[c(s), t]. \tag{2.3}$$

Hence (2.2) yields the fundamental functional equation

$$f(c, s + t) = \operatorname*{Max}_{z\,[0,\,s]} \left\{\int_0^s F(x, z)\, du + f[c(s), t]\right\}. \tag{2.4}$$

The notation $z[0, s]$ indicates that the maximization is over all $z(u)$ taken over

$0 \leq u \leq s$. As s shrinks to 0, the choice of a function reduces to a choice of $z(0)$, a quantity which we shall call w for typographical convenience. A straightforward calculation yields

$$f_t = \operatorname{Max}_w [F(c, w) + G(c, w)f_c] \tag{2.5}$$

as the limiting form of (2.4) as $s \to 0$. The maximizing w is obtained by equating the partial derivative with respect to w to zero, $0 = F_w + G_w f_c$.[12] Combining this with (2.5), we obtain the system of equations

$$\begin{aligned} f_t &= (FG_w - GF_w)/G_w = K(c, w), \\ f_c &= -F_w/G_w = L(c, w). \end{aligned} \tag{2.6}$$

To derive a partial differential equation for w, we employ the identity $(f_t)_c = (f_c)_t$. The result is

$$K_c + K_w w_c = L_w w_t. \tag{2.7}$$

As is well known,[13] the general solution of this equation may be obtained from the general solution of a system of first-order ordinary differential equations. Since c is known for $t = 0$, to complete the solution we require only the value of w at $t = 0$. For small t we have

$$\int_0^t F(x, z)\, du = F(c, w)t + 0(t^2). \tag{2.8}$$

Hence w at $t = 0$ is determined by $F_w = 0$.

Taking $G(x, z) = z$, we are reduced to the familiar problem of maximizing $\int_0^t F(x, x')\, du$. It is readily verified that the characteristic equations obtained from (2.7) are equivalent to the usual Euler equation.

The general problem of maximizing $\int_0^t F(x_1, x_2, \ldots, x_n; z_1, z_2, \ldots, z_m)du$ subject to the constraints $dx_i/dt = G_i(x, z)$, $x_i(0) = c_i$ may be attacked in the same fashion.

3. *The Eigenvalue Problem.*—The eigenvalue problem posed in §1 is equivalent to that of determining the successive minima of $\int_0^1 x'^2\, du$ subject to the constraints $\int_0^1 \phi x^2\, du$, $x(0) = x(1) = 0$. To treat the problem using the functional-equation approach, we imbed it within the more general problem of determining the minima of $J(x) = \int_a^{a+t} x'^2\, du$ subject to the constraints

$$\begin{aligned} &(a) \quad x(a) = x(a + t) = 0, \\ &(b) \quad \int_a^{a+t} \phi(u)x^2\, du + k \int_a^{a+t} \phi(u)\,(a + t - u)x\, du = 1. \end{aligned} \tag{3.1}$$

Let $\operatorname{Min} J(u) = f(a, k, t)$. Then, if, as above, we write

$$f(a, k, s + t) = \int_a^{a+s} x'^2\, du + \int_{a+s}^{a+s+t} x'^2\, du, \tag{3.2}$$

we cannot derive a functional equation immediately, since $x(a + s)$ is not necessarily zero. To simplify the analysis, let us pass directly to the derivation of the partial differential equation, assuming that s is an infinitesimal. It is sufficient to consider only terms which are of zeroth or first order in s. Since $x(a + s) = sx'(a) + \ldots$, we set $y(u) = x(u) - sx'(a)\,(a + s + t - u)/t$. Then $y(a + s) =$

$y(a + s + t) = 0$, to the order of our approximation, and $x'(u) = y'(u) - sx'(a)/t$. Thus, to terms in s^2,

$$\int_{a+s}^{a+s+t} x'^2\,du = \int_{a+s}^{a+s+t} y'^2\,du.$$

The constraint of (3.1b) becomes

$$\int_{a+s}^{a+s+t} \phi(u)y^2\,du + \left[k + \frac{2sx'(a)}{t}\right]\int_{a+s}^{a+s+t} \phi(u)y(u)(a + s + t - u)\,du = 1 - s\psi[k, x'(a), t] + 0(s^2), \quad (3.3)$$

where

$$\psi[k, x'(a), t] = \frac{kx'(a)}{t}\int_a^{a+t} \phi(u)\,(a + t - u)^2\,du. \quad (3.4)$$

Making a change of variable $y = (1 - (s\psi/2))w$, to reduce the constant term in (3.3) to 1, the constraint takes the form

$$\int_{a+s}^{a+s+t} \phi(u)w^2\,du + \left\{k + s\left[\frac{k\psi}{2} + \frac{2x'(a)}{t}\right]\right\}\int_{a+s}^{a+s+t} \phi(u)w(u)\,(a + s + t - u)\,du = 1 + 0(s^2). \quad (3.5)$$

The equation in (3.2) becomes

$$f(a, k, s + t) = \min_{x'[a,\,a+s]} \left\{\int_a^{a+s} x'^2\,du + (1 - s\psi)f\left[a + s, k + s\left(\frac{k\psi}{2} + \frac{2x'[a]}{t}\right), t\right]\right\} + 0(s^2). \quad (3.6)$$

Letting $s \to 0$, this yields the partial differential equation

$$f_t = \operatorname*{Min}_{x'(a)} \left\{x'(a)^2 + f_a + \left[\frac{k\psi}{2} + \frac{2x'(a)}{t}\right]f_k - \psi f\right\}. \quad (3.7)$$

Since $\psi = x'(a)\Phi$ is independent of $x'(a)$, the unique minimum occurs when

$$2x'(a) + \left(\frac{k\Phi}{2} + \frac{2}{t}\right)f_k - \Phi f = 0. \quad (3.8)$$

Eliminating $x'(a)$ by means of this last relation, we obtain a nonlinear partial differential equation for f.

The behavior of f for small t may be determined by solving the problem of minimizing $\int_a^{a+t} x'^2\,du$ subject to the constraints

$$x(a) = x(a + t) = 0, \quad (3.9)$$

$$\phi(a)\int_a^{a+t} x^2\,du + k\int_a^{a+t} \phi(u)\,(a + t - u)x\,du = 1.$$

This problem, in turn, may be treated by the partial differential equation approach, using the fact that the solution for $k = 0$ is immediate.

[1] R. Bellman, "On the Theory of Dynamic Programming," PROC. NATL. ACAD. SCI., **38**, 716–719 (1952).

[2] R. Bellman, "On Bottleneck Problems and Dynamic Programming," *ibid.*, **39**, 947–951 (1953).

[3] R. Bellman, "On Computational Problems in the Theory of Dynamic Programming," *Symposium on Applied Mathematics*, Santa Monica, 1953.

[4] R. Bellman, "Some Functional Equations in the Theory of Dynamic Programming," PROC. NATL. ACAD. SCI. (to appear).

[5] R. Bellman, "On Limiting Theorems for Non-commutative Processes," *Duke Math. J.* (to appear).

[6] R. Bellman, "A Problem in the Theory of Dynamic Programming," *Econometrica* (to appear).

[7] R. Bellman, "An Optimal Inventory Equation," sub. to *Pac. J. Math.*

[8] R. Bellman, I. Glicksberg, and O. Gross, "On Some Variational Problems in the Theory of Dynamic Programming," PROC. NATL. ACAD. SCI., **39**, 298–301 (1953).

[9] R. Bellman, T. E. Harris, and H. N. Shapiro, *Some Functional Equations Occurring in the Theory of Decision Processes*, Rand Papers, No. 382.

[10] R. Bellman and S. Lehman, "On the Continuous Gold-Mining Equation," PROC. NATL. ACAD. SCI. (to appear).

[11] R. Bellman and O. Gross, "Some Combinatorial Problems Arising in the Theory of Multi-Stage Production Processes" (to appear).

[12] A similar formalism has been applied by R. Isaacs in his theory of differential games, i.e., systems of differential equations of the form $dx_i/dt = \operatorname{Min}_{w} \operatorname{Max}_{z} F_i(x, z, w)$.

[13] R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, New York: Interscience Publishers, 1943.

A CHARACTERIZATION OF TAME CURVES IN THREE-SPACE

BY O. G. HARROLD, JR., H. C. GRIFFITH,* AND E. E. POSEY

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TENNESSEE

Communicated by G. T. W. Whyburn, March 4, 1954

1. *Introduction.*—A subset K of 3-space has been called tamely imbedded if there is a homeomorphism on space throwing K onto a polyhedron. If K is a simple arc or simple closed curve (i.e., a 1-manifold) that is contained in a plane, it follows quickly from classical results of Schoenflies that K is tamely imbedded. If the 1-manifold is not a subset of a plane, it is necessary to distinguish whether K is an arc or a closed curve. If K is a polygonal arc, it is already tame and also has the property of being equivalent to an interval under a semilinear map on 3-space. If K is a polygonal closed curve, it is already tame, but the intermediate step of mapping K into a plane is neither advantageous nor possible if K is knotted.

About thirty years ago an example was given of an arc K which is not equivalent to a subset of a plane under any homeomorphism on 3-space. This example and others closely related to it, including some simple closed curves, were studied by Antoine,[1] Alexander,[2] and later by Fox and Artin[3] and serve to emphasize the pathological difficulties that may occur. The nature of these examples suggests that the difficulties are of a local nature as regards the manner the 1-manifold is imbedded in 3-space. This point of view is borne out by recent results of R. H. Bing[4] and E. E. Moise,[5] who prove that a locally tamely imbedded set is tamely imbedded.

The purpose of the present paper is to characterize, by means of positional invariants, those 1-manifolds which are locally tamely imbedded. By the above-mentioned result, this implies a characterization of the tamely imbedded 1-mani-

folds. If we restrict attention solely to simple closed curves, the purpose may be put thus: to distinguish the curves which determine classical knot-types from those that are pathological.

2. *Definitions.*—The symbols $\cap$ and $\cup$ have their usual set-theoretic meaning, $A \setminus B$ means the collection of points in A but not in B, and $S(A, \epsilon)$ means the collection of points whose distances to some point of A is less than ϵ. If B is a homogeneous polyhedron, $|\partial B|$ will mean the point set carrying the chain ∂B. Three-space is denoted by R^3.

The set $X \subset R^3$ is *locally polyhedral* at $p \in R^3$, provided that there is some neighborhood N of p in R^3 whose intersection with X is null or a finite polyhedron. The set X is called *locally polyhedral modulo* Y (mod Y) if it is locally polyhedral at each point of $R^3 \setminus Y$.

A homeomorphism h of R^3 on R^3 is called *semilinear* if and only if there are subdivisions σ and τ of R^3 into simplices such that every simplex of σ is mapped affinely on some simplex of τ.

Let J be an arc or simple closed curve. Then J is said to have *property* $\mathcal{P}$ provided that for each $\epsilon > 0$ and point $x \in J$ there is a topological 2-sphere $K \subset S(x, \epsilon)$ which is locally polyhedral at points of $K \setminus J$, contains x in its interior, and meets J in a set whose cardinal is the (Menger-Urysohn) order of x in J.[6]

An arc (simple closed curve) J is said to have *property* $\mathcal{Q}$, provided that there is a topological closed 2-cell G (a disk) with $J \subset |\partial G|$ ($J = |\partial G|$) which is locally polyhedral modulo J.

An arc (simple closed curve) J is said to have *property* $\mathcal{Q}$ at x, provided that there is a closed 2-cell G such that (i) G is locally polyhedral mod J; (ii) $G \cap J$ is an arc (in $|\partial G|$); and (iii) $G \cap J$ is the closure of a neighborhood of x relative to J.

Since knotted polygonal curves exist, J may have property $\mathcal{Q}$ at every point and still fail to have property $\mathcal{Q}$. Examples showing that neither property $\mathcal{P}$ nor property $\mathcal{Q}$ alone is sufficient to guarantee tame imbedding are easily constructed. Fox and Artin's example 1.4[3] is an arc having property $\mathcal{P}$ but is wildly imbedded. Joining the end-points of this arc together by a segment meeting it only at its end-points provides a wild curve with property $\mathcal{P}$. Example 2.1 of the same paper is a curve bounding a 2-cell, which may be taken locally polyhedral save at two points of its boundary and which has a boundary that is wildly imbedded.

3. *The Characterization.*—It is first shown that the arc problem depends on the closed-curve problem. That is, if an arc J has the properties $\mathcal{P}$ and $\mathcal{Q}$, there is a simple closed curve $K \supset J$ that has properties $\mathcal{P}$ and $\mathcal{Q}$. The following theorems are then proved.

I. *If there is a homeomorphism of R^3 onto R^3 carrying J into a subset of a plane, then J has properties $\mathcal{P}$ and $\mathcal{Q}$.*

It is not obvious that the properties $\mathcal{P}$ and $\mathcal{Q}$ defined above are positional invariants. The arguments given to prove Theorem I establish this fact.

II. *If J is a simple closed curve in R^3 having properties $\mathcal{P}$ and $\mathcal{Q}$, then J is tame.*

The method of proof of this theorem may be briefly indicated. Let $\mathfrak{J}$ denote the unit circle in the xy-plane, $\mathfrak{D}$ the disk in the xy-plane bounded by $\mathfrak{J}$, and $\mathfrak{T}_i$ the torus bounding the region $S(\mathfrak{J}, 1/i + 1)$, $i = 1, 2, \ldots$. Let $\{\mathfrak{a}_i\}$ be a countable dense subset of $\mathfrak{J}$ and let $\mathfrak{D}_i$ be the disk which is the intersection of the closure of $S(\mathfrak{a}_i, 1/1 + i)$ and the plane determined by the z-axis and $\mathfrak{a}_i$, $i = 1, 2, \ldots$.

Then, since J has properties $\mathcal{P}$ and $\mathcal{Q}$, sets G, T_i, and D_i may be constructed so that the sets $G \cup T_i \cup D_i$ and $\mathfrak{G} \cup \mathfrak{T}_i \cup \mathfrak{D}_i$ are homeomorphic.

The extension of this homeomorphism to a self-homeomorphism of R^3 which maps J onto $\mathfrak{J}$ is then made. In obtaining this extension strong use is made of the result due to Alexander[7] that one of the complementary domains of a polyhedral torus in spherical 3-space is homeomorphic to a complementary domain of a standard torus.

It is known[8] that if S_1, S_1' and S_2, S_2' are two pairs of polyhedral spheres in R^3, with S_i contained in the interior of S_i' (Int S_i'), then there is a homeomorphism of the closure of the set (Int S_1') $\cap$ (Ext S_1) onto the closure of (Int S_2') $\cap$ (Ext S_2). It is evident that quite a diverse collection of types of regions are bounded by pairs of disjoint polyhedral tori, but an extension of the above-mentioned result to suitably restricted pairs of "concentric" polyhedral tori is made. This provides a chief tool for the proof of the above theorem.

III. *If the 1-manifold J has properties $\mathcal{P}$ and $\mathcal{Q}$ at each point, then J is locally tame and conversely.*

* This work was begun jointly during the tenure of a National Science Foundation Fellowship by Griffith.

[1] L. Antoine, "Sur la possibilité d'entendre l'homéomorphisme de deux figures a leur voisinage," *Compt. rend. Acad. Sci.*, **171**, 661–663 (1920).

[2] J. W. Alexander, "An Example of a Simply Connected Surface Bounding a Region Which Is Not Simply Connected," PROC. NATL. ACAD. SCI., **10**, 8–10 (1924).

[3] R. Fox and E. Artin, "Some Wild Cells and Spheres in 3-Dimensional Space," *Ann. Math.*, **49**, 979–990 (1948).

[4] R. H. Bing, "Locally Tame Sets Are Tame," *Ann. Math.*, **59**, 145–148 (1954).

[5] E. E. Moise, "Affine Structures in 3-Manifolds. VIII," *Ann. Math.*, **59**, 159–170 (1954).

[6] O. G. Harrold, Jr., "The Enclosing of Simple Arcs by Simple Closed Surfaces in Three-Space," *Bull. Amer. Math. Soc.*, **59**, 369 (1953). To appear in the *Duke Mathematical Journal*.

[7] J. W. Alexander, "On the Sub-division of Space by a Polyhedron," PROC. NATL. ACAD. SCI., **10**, 6–8 (1924).

[8] W. Graeub, "Die semi-lineare Abbildungen," *Sitzungsb. Heidelberger Akad. Wissensch.*, Part N, pp. 205–272 (1950).

CONFORMAL GEOMETRY AND ELEMENTARY PARTICLES

BY R. INGRAHAM

INSTITUTE FOR ADVANCED STUDY, PRINCETON, N.J.

Communicated by Oswald Veblen, February 25, 1954

If the conformal replaced the Lorentz geometry as the geometric base of physical field theories, what new physical ideas could be expected? On the purely kinematical level, that is, before even the assumption of a particular set of (classical or quantized) field laws, much can be said. This note summarizes the main finding of a paper to appear shortly. The space in question is the 5-dimensional one of all spheres in R_4 (flat 4-space of sig. $[+++-]$); the fundamental invariant, the angle θ, under which two spheres intersect. This is an essential difference from former attempts to apply the conformal geometry to physics, which used only the 4-dimensional space of points (null spheres) as the domain of the physical fields. Our re-

sults may be summed up briefly by saying that a new concept of elementary particle results. From being an entity superimposed on the basic geometry, as is the present view, it becomes integrated into the basic geometry, from which it is logically inseparable. Any geometry built on this geometry must have the following kinematical properties.

1. The spheres are labeled by the physical co-ordinates $(\vec{r}, t, \omega)$ (Dim L, T, M^{-1} resp.) when considered as the spheres

$$(\vec{\xi} - \vec{r})^2 - c^2(\tau - t)^2 = \epsilon \left(\frac{\hbar\omega}{c}\right)^2 \qquad (\epsilon = \pm 1) \quad (1)$$

in space-time ("q-space"), $(\vec{\xi}, \tau)$ Lorentz co-ordinates, and by $(\vec{p}, E, m)$ (Dim MLT^{-1}, ML^2T^{-2}, M resp.) when considered as the spheres

$$(\vec{\pi} - \vec{p})^2 - 1/c^2(\eta - E)^2 = \epsilon(mc)^2 \qquad (2)$$

in momentum-energy space ("p-space"), $(\vec{\pi}, \eta)$ Lorentz co-ordinates. In these ("natural") co-ordinates the angle metric has the forms, resp.,

$$d\theta^2 = \epsilon \left(\frac{\hbar\omega}{c}\right)^{-2} (dr^2 - c^2\, dt^2 - \epsilon\hbar^2/c^2\, d\omega^2),$$
$$d\theta^2 = \epsilon(mc)^{-2}\, (dp^2 - 1/c^2\, dE^2 - \epsilon c^2\, dm^2). \qquad (3)$$

The "sphere" (hyperboloid as real locus) (1) represents a finite (i.e., nonlocalized) test particle of *size* $\hbar\omega/c$ "at" the event $(\vec{r}, t)$ in q-space. Similarly, (2) represents a finite test particle of rest mass m "at" the momentum-energy $(\vec{p}, E)$ in p-space. The nonlocalizations in q- and p-space are supposed to vary directly with $\hbar\omega/c$ and mc, resp.[1] Fields will thus be defined on these *particle states* (thing labeled by $[\vec{r}, t, \omega]$ or $[\vec{p}, E, m]$) rather than on the events or 4-momenta alone.

2. The group of conformal transformations in R_4 (proper Lorentz transformations, translations, constant dilations, inversions in spheres) is just the 15-parameter group of motions (metric-preserving transformations) of the Riemannian 5-space (of constant unit curvature) with metric (3). These motions connect the physically equivalent frames of the theory. A special motion

$$p^\alpha = -\hbar x^\alpha/[x]^2, \quad \text{inverse } x^\alpha = -\hbar p^\alpha/[p]^2, \quad (\alpha = 1, \ldots, 5)$$
$$[x]^2 = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2 - \epsilon\,(x^5)^2, \qquad (4)$$

where $x^1 = x$, etc., $x^4 = ct$, $x^5 = \hbar\omega/c$; $p^1 = p_x$, etc., $p^4 = E/c$, $p^5 = cm$, exists, which takes one from q-space (dim $x^\alpha = L$) to the associated p-space (dim $p^\alpha = MLT^{-1}$) and vice versa. Hence any conformal theory exhibits an automatic reciprocity[2] between q- and p-space having the two aspects: (a) any equation in q-space gives another true equation in p-space on making the substitution $x^\alpha \to p^\alpha$ in it, and conversely; (b) any solution for the fields relative to a q-frame tensorially transformed under (4) gives the same solution relative to the associated p-frame,

and conversely. Moreover, the conjugate co-ordinates p^α and x^α of a particle state satisfy by (4) the uncertainty-relation type equation,

$$[px] = \vec{p} \cdot \vec{r} - Et - \epsilon\, \hbar\omega m = -\hbar, \tag{5}$$

so that the nonlocalizations of the test particle in p- and q-space are in a reciprocal relation, Planck's $\hbar$ being a measure of the intrinsic correlation. The subgroup of motions involving relative motion of observers represents *uniform* relative acceleration of Lorentz frames as found by Haantjes[3] (Lorentz group $\longleftrightarrow$ zero accelerations). Test particle size and mass are invariant under the Lorentz group but noninvariant under the accelerative motions, being renormalized nonuniformly (depending on x^α or p^α). The group of motions can be built up by adjoining to the subgroup preserving the physical dimension the single transformation (4) which switches q- and p-space.

3. The motion of test particles is described by geodesics in sphere space—i.e., certain 1-dimensional families of particle states. In our case (force-free) they can be written

$$\begin{aligned} &\vec{r} - \vec{r}_0 = \vec{v}\,(t - t_0), \quad \hbar\omega/c = \left|1 - v^2/c^2\right|^{1/2} c(t - t_0), \\ &\vec{p} - \vec{p}_0 = \vec{u}\left(\frac{E - E_0}{c^2}\right), \quad mc = \left|1 - u^2/c^2\right|^{1/2}\left(\frac{E - E_0}{c}\right), \end{aligned} \tag{6}$$

the null geodesics; and

$$\begin{aligned} &\vec{r} - \vec{r}_0 = \vec{v}(t - t_0), \quad \hbar\omega/c = \left|1 - v^2/c^2\right|^{1/2} \times \left[c^2(t - t_0)^2 + \left(\frac{\lambda_0}{1 - v^2/c^2}\right)^2\right]^{1/2}, \\ &\vec{p} - \vec{p}_0 = \vec{u}\left(\frac{E - E_0}{c^2}\right), \quad mc = \left|1 - u^2/c^2\right|^{1/2} \times \left[\left(\frac{E - E_0}{c}\right)^2 + \left(\frac{\hbar/l_0}{1 - u^2/c^2}\right)^2\right]^{1/2}, \end{aligned} \tag{7}$$

the non-null geodesics. $\vec{v}, \vec{r}_0, t_0, \lambda_0, \vec{u}, \vec{p}_0, E_0, l_0$ are (dimensional) integration constants, and the condition $\mathrm{Sgn}(1 - v^2/c^2) = \mathrm{Sgn}(1 - u^2/c^2) = -\epsilon$ is imposed. ($\epsilon = +1$ seems to be excluded, as giving spacelike motion here and a spacelike relation between the spheres [1] and [2] and their centers.) These are free elementary particle paths. There are good reasons for thinking that the null geodesics represent photon and electron propagation. $v = c$ or $u = c$ (conformal covariant statement) defines the photon; it follows that $\omega = m = 0$. The increase of $\hbar\omega/c$ with time and mc with energy shows that the free particle tends to dislocalize in q- and p-space with increasing time, energy resp. Born conjugating[2] the q-paths show that $\vec{p}$, E, and m are (asymptotically) constant in time and related to one another with $\vec{u} = \vec{v}$ as expected from special relativity dynamics. *This result here follows from these motion equations.* Free elementary particles *cannot* be accelerated by the uniform accelerative changes of frame, since the uniform-velocity form of (6)

and (7) holds for all our equivalent frames. This contradicts relativity dynamics and suggests a Michelson-Morley type experiment capable in principle of choosing between the conformal and Lorentz geometries for physics.

[1] The exact observational significance of these hyperboloids is not yet clear. Concerning this model of nonlocalized particles, see H. Freistadt, *Phys. Rev.*, **92**, 1015 (1953).

[2] This reciprocity may be called *Born reciprocity*, and the special motion (4), *Born conjugation*, after the principal exponent of this sort of idea in modern physics.

[3] J. Haantjes, *Proc. Kon. Ned. Akad. Wet.*, **43**, 1288 (1940); cf. also E. Hill, *Phys. Rev.*, **73**, 143 (1947).

APPROXIMATION BY BOUNDARY VALUES OF ANALYTIC FUNCTIONS

By Jacqueline Penez

UNIVERSITY OF MINNESOTA

Communicated by Marston Morse, February 17, 1954

F. Riesz[1] has investigated the problem of minimizing the integral

$$\int_{|z|=1} |f(z) - p(z)|\,|dz|,$$

where $f(z)$ is a prescribed polynomial in $1/z$ and $p(z)$ ranges over a class of boundary functions of functions analytic for $|z| < 1$. Extremal problems of a similar type have been considered by Kakeya,[2] Golusin,[3] Doob,[4] Garabedian and Schiffer,[5] MacIntyre and Rogosinski.[6] In this paper we will discuss the integral

$$\int_C |f(z) - p(z)|^k |dz|$$

for $k > 1$, where C is the rectifiable boundary of a finitely multiply connected domain D; $f(z)$ is a complex-valued function whose kth power is Lebesgue integrable on C; and $p(z)$ ranges over the boundary functions of functions analytic in D.

More precisely, we let D be a finite region of the z-plane bounded by a system of curves C, consisting of an outer rectifiable Jordan curve C_0 and v rectifiable Jordan curves $C_1, \ldots, C_v$ lying within C_0 and having no points in common with it or with each other. We let $z(s)$ be the parametric representation of C in terms of arc length along C. By $L^k(C)$, where $k \geq 1$, we denote the class of all complex-valued functions $f(z)$ defined on C such that

$$(\|f\|_k)^k = \int_C |f(z)|^k\, ds$$

exists finite in the Lebesgue sense. Since $\|f - g\|_k = 0$ if and only if $f = g$ almost everywhere on C, we replace the original space $L^k(C)$ by the space of equivalence classes determined by the relation $\|f - g\|_k = 0$ and henceforth refer to the latter class as $L^k(C)$. The class $B^k(C)$ will denote the subclass of functions in $L^k(C)$ which satisfy the following conditions:

$$\int_C f(z) z^n\, dz = 0 \text{ for } n = 0, 1, 2, \ldots, \tag{1}$$

$$\int_C f(z)(z - a_i)^{-n}\, dz = 0 \text{ for } n = 1, 2, \ldots; i = 1, \ldots, v, \tag{2}$$

where $a_i \in \text{Int } (C_i)$. Another subclass of $L^k(C)$ to be considered is

$$A^k(C) = \overline{\{f | C | f \text{ analytic in } \bar{D}\}},$$

where the closure is taken in the space $L^k(C)$. We note that

$$A^k(C) \subseteq B^k(C) \text{ for } k \geq 1.$$

Definition of Smirnoff Domains.—Case (i): Let D be the interior of a rectifiable Jordan curve. If $z = \varphi(t)$ maps $|t| < 1$ conformally onto D, then, by a theorem of F. and M. Riesz, $\log |\varphi'(t)| \in H_1$ ($|t| < 1$), and consequently $\log |\varphi'(t)|$ has a Poisson-Stieltjes integral representation. If, moreover, $\log |\varphi'(t)|$ has a Poisson-Lebesgue integral representation, then D will be said to be of Smirnoff type.[7, 8]

Case (ii): Let D be the exterior of a rectifiable Jordan curve C and $a \in \text{Int } (C)$, $\{z | |z - a| \leq r\} \subseteq \text{Int } (C)$. In this case, D will be said to be of Smirnoff type if D' is of Smirnoff type, where D' is the image of D under $z = \psi(w) = (r/w) + a$.

Case (iii): Let D be a finitely multiply connected domain with a boundary C as described above. Then D is of Smirnoff type if $\text{Int}(C_0)$ and $\text{Ext}(C_i)$ for $i = 1, \ldots, v$ are domains of Smirnoff type. We let S denote the class of domains of Smirnoff type.

To $f(z) \in B^k(C)$ corresponds a function $F(z)$ analytic in D, whose nontangential boundary values coincide with $f(z)$ (almost everywhere) on C. Hence $B^k(C)$ is a class of boundary functions of functions analytic in D. In particular, if D is simply connected and $D \in S$, then $B[H_k(D)] = B^k(C)$ for $k \geq 1$ where $B[H_k(D)]$ denotes the class of boundary functions of functions in $H_k(D)$, the usual Hardy classes.[9, 10, 11]

THEOREM 1. If $D \in S$, $f(z) \in B^k(C)$ and $g(z) \in B^q(C)$, where $1/k + 1/q = 1$, then

$$\int_C f(z)g(z)\,dz = 0.$$

COROLLARY 1. If $D \in S$, then $A^k(C) = B^k(C)$ for $k > 1$. This follows from Theorem 1 by the Hahn-Banach theorem.

COROLLARY 2. If $D \in S$ and D is simply connected, then $B[H_k(D)] = B^k(C) = A^k(C)$ for $k \geq 1$.

Now we proceed to the extremal problem itself. Both classes $B^k(C)$ and $A^k(C)$ are closed linear subspaces of $L^k(C)$, and the spaces $L^k(C)$ for $k > 1$ are uniformly convex Banach spaces.[12] From these facts, we deduce Theorem 2.

THEOREM 2 (Existence and Uniqueness Theorems).

Let $f(z) \in L^k(C)$, where $k > 1$. Then there exists a unique function $g(z) \in B^k(C)$ such that

$$\|f - g\|_k = \inf_{h \in B^k(C)} \|f - h\|_k.$$

Similarly, there exists a unique function $g(z) \in A^k(C)$ such that

$$\|f - g\|_k = \inf_{h \in A^k(C)} \|f - h\|_k.$$

We shall call the function $g(z)$ which minimizes $\|f - h\|_k$ over $B^k(C)$ [or $A^k(C)$] the *minimizing* function, and the difference $f(z) - g(z)$ a *minimal* function.

THEOREM 3 (Characterization of the Minimal Functions).

Let $0 \not\equiv f(z) \in L^k(C)$, where $k > 1$. Then $f(z)$ is minimal over $A^k(C)$ [or $B^k(C)$] if and only if $|f|^k/z'(s)f$ coincides (almost everywhere) on C with a function $F(z) \in B^q(C)$

[or $A^q(C)$], where $1/k + 1/q = 1$. We note that minimal functions can vanish at most on a set of measure zero, since a function in $B^q(C)$ where $q \geq 1$ vanishes at most on a set of measure zero.

This extremal problem is related to another extremal problem, namely, that of finding the norms of linear functionals on the classes $B^k(C)$ and $A^k(C)$. This maximum-minimum duality in extremal problems has been studied by MacIntyre and Rogosinski[6] and Shapiro,[13] and our results overlap with theirs for the case of the unit circle. In this connection, we have the following theorem:

THEOREM 4. For each $\omega(z) \in L^q(C)(q > 1)$, where $1/k + 1/q = 1$, the function L defined for $f(z) \in B^k(C)$ [or $A^k(C)$] by

$$L(f) = \int_C f(z)\omega(z)\, dz$$

is a linear functional on $B^k(C)$ [or $A^k(C)$] and

$$\|L\|_{B^k} = \max_{\substack{\|f\|_k \leq 1 \\ f \in B^k(C)}} |L(f)| = \min_{g \in Aq} \|\omega - g\|_q,$$

$$\text{or } \|L\|_{A^k} = \max_{\substack{\|f\|_k \leq 1 \\ f \in A^k}} |L(f)| = \min_{g \in Bq} \|\omega - g\|_q.$$

In fact,

$$\|L\|_{B^k(\text{or } A^k)} = \left| \frac{L(F)}{\|F\|_k} \right| = \|\omega - g_0\|_q,$$

where $g_0(z)$ is the minimizing function over $A^q(C)$ [or $B^q(C)$] of $\omega(z)$ and $F(z) = |\omega(z) - g_0(z)|^q/z'(s)\,[\omega(z) - g_0(z)]$. Conversely, if L is a linear functional on $B^k(C)$ [or $A^k(C)$], then there exists $\omega(z) \in L^q(C)$ such that $L(f) = \int_C f(z)\omega(z)\, dz$ and $\|L\|_{Bk}$ (or $\|L\|_{Ak}$) is as above.

We remark that a linear functional on $B^k(C)$ [or $A^k(C)$] is represented in terms of a minimal function over $A^q(C)$ [or $B^q(C)$], since

$$\int_C f(z)\omega(z)\, dz = \int_C f(z)\,[\omega(z) - g_0(z)]\, dz,$$

when $f(z) \in B^k(C)$ [or $A^k(C)$] and $g_0(z) \in A^q(C)$ [or $B^q(C)$].

Let us now discuss briefly an application for the case where D is a simply connected domain of Smirnoff type. We remark that, from the work of J. L. Doob,[4] it is easy to deduce the existence, uniqueness, and characterization for the case $k = 1$. The problem of finding minimizing functions for given functions in $L^k(C)$ $(k \geq 1)$ reduces to that of solving the problem for the unit circle by the following theorem:

THEOREM 5. Let D be a simply connected region bounded by a simple rectifiable Jordan curve C and $z = \varphi(t)$ map $|t| < 1$ conformally onto Int(C) and $t = \Phi(z) = \varphi^{-1}(z)$. Let γ be $|t| = 1$, $D \in S$, and $f(z) \in L^k(C)$, $k \geq 1$.

(i) If $f(z)$ is minimal over $B^k(C)$ $(k \geq 1)$, then $f[\varphi(t)]\,[\varphi'(t)]^{1/k}$ is minimal over $B^k(\gamma)$ and conversely.

(ii) If $G(t)$ is the minimizing function for $f[\varphi(t)]\,[\varphi'(t)]^{1/k}$ over $B^k(\gamma)$, then $G[\Phi(z)]\,[\Phi'(z)]^{1/k}$ is the minimizing function for $f(z)$ over $B^k(C)$.

Application.—Let D be as in Theorem 5 and $\alpha \in D$. The minimal function for $1/(z - \alpha)$ over $B^k(C)$ for $k \geq 1$ is

$$\frac{[\Phi'(\alpha)]^{1/q}\ [\Phi'(z)]^{1/k}}{\Phi(z)\ -\ \Phi(\alpha)}\left[\frac{1\ -\ |\Phi(\alpha)|^2}{1\ -\ \overline{\Phi(\alpha)}\ \Phi(z)}\right]^{1/k\ -\ 1/q}$$

If we choose $\Phi(z)$ such that $\Phi(\alpha) = 0$, then the minimal function for $1/(z - \alpha)$ reduces to $[\Phi'(\alpha)]^{1/q}[\Phi'(z)]^{1/k}/\Phi(z)$, which we will denote by $h(\alpha, z)$ for simplicity. We note that theoretically the problem of finding the mapping function is equivalent to that of finding the minimal function for $1/(z - \alpha)$. Since $h(\alpha, z)$ is minimal, then $|h(\alpha, z)|^k/z'(s)h(\alpha, z)$ coincides (almost everywhere) on C with a function $p(z) \in B^q(C)$. As usual, we define $p(z)$ for $z \in \text{Int}(C)$ by Cauchy's integral and $p(z) \in H_q(D)$.

The function L, defined for $f(z) \in B^q(C)$ $(q > 1)$ by

$$L(f) = \int_C \frac{f(z)}{z - \alpha}\, dz,$$

is a linear functional and Theorem 4 implies that

$$|f(\alpha)| \leq \frac{1}{2\pi}\,\|f\|_q\,\|h(\alpha, z)\|_k = \frac{1}{(2\pi)^{1/q}}\,\|f\|_q\,|\Phi'(\alpha)|^{1/q},$$

and equality is obtained when and only when $f(z) = p(z)$.

One needs only to vary the kernel $\omega(z)$ in Theorem 4 to obtain other interesting results.

[1] F. Riesz, "Über Potenzreihen mit vorgeschriebenen Anfangsgliedern," *Acta Math.*, **42**, 145–171 (1920).

[2] S. Kakeya, "General Mean Modulus of Analytic Functions," *Proc. Phys.-Math. Soc. Japan*, **3**, 48–58 (1921).

[3] G. M. Golusin, "On the Problem of Caratheodory-Fejer and Similar Problems," *Mat. Sbornik*, **18**, 213–226 (1946).

[4] J. L. Doob, "A Minimum Problem in the Theory of Analytic Functions," *Duke Math. J.*, **8**, 413–24 (1941).

[5] P. R. Garabedian and M. Schiffer, "On Existence Theorems of Potential Theory and Conformal Mapping," *Ann. Math.*, **52**, 164–187 (1950).

[6] A. J. MacIntyre and W. W. Rogosinski, "Extremum Problems in the Theory of Analytic Functions," *Acta Math.*, **82**, 275–325 (1950).

[7] V. J. Smirnoff, "Sur les formules de Cauchy et de Green et quelques problèmes qui s'y rattachent," *Bull. Acad. Sci. U.R.S.S.*, Sér. VII, 337–371 (1932).

[8] M. V. Keldysch and M. Lavrentieff, "Sur la représentation conforme des domaines limités par des courbes rectifiables," *Ann. Sci. École Normale Supérieure*, **54**, 1–38 (1937).

[9] F. and M. Riesz, "Über die Randwerte einer analytischen Funktion," *Compt. Rend. Quatrième Cong. des Mathématiciens Scandinaves à Stockholm*, pp. 27–44 (1916).

[10] F. Riesz, "Über die Randwerte einer analytischen Funktion," *Math. Zeitschr.*, **18**, 87–95 (1923).

[11] I. Privaloff, *Boundary Properties of Analytic Functions*. 2d ed. Moscow, 1950 (in Russian).

[12] J. A. Clarkson, "Uniformly Convex Spaces," *Trans. A.M.S.*, **40**, 396–414 (1936).

[13] H. S. Shapiro, "Maximum-minimum Duality in Extremal Problems," *Bull. A.M.S.*, **58**, 570 (1952).

*A NOTE ON THE ABSTRACT CAUCHY PROBLEM**

By R. S. Phillips†

YALE UNIVERSITY AND THE UNIVERSITY OF SOUTHERN CALIFORNIA

Communicated by E. Hille, February 6, 1954

1. The notion of an abstract Cauchy problem has recently been introduced by E. Hille.[1] In the present note we show that if the Cauchy problem is solvable in a restricted sense with respect to a given closed linear operator U for all initial values in the domain of U, then U is the infinitesimal generator of a semigroup yielding the same solutions to the Cauchy problem. Under our assumptions, the existence of a unique solution to the Cauchy problem for all initial values in the domain of U is in itself sufficient to force the solutions to vary continuously with the initial data.

2. We begin with a summary of semigroup theory.[2] A family of linear bounded operators $[S(t); t > 0]$ on a complex Banach space $\mathfrak{X}$ to itself is called a semigroup if $S(t_1 + t_2) = S(t_1)S(t_2)$ for $t_1, t_2 > 0$. $S(t)$ will be assumed to be continuous in the strong-operator topology for $t > 0$. In this case $w_0 \equiv \lim_{t\to\infty} t^{-1} \log \|S(t)\| < \infty$. The semigroup is said to be of class (C_0) if $\lim_{t\to 0+} S(t)x = x$ for each $x \in \mathfrak{X}$; it then follows that, given $w > w_0$, there exists an $M > 0$ such that $\|S(t)\| \leqq M \exp(wt)$, $t > 0$. A semigroup is said to be of class $(0, A)$ if $\int_0^1 \|S(t)x\|\, dt < \infty$ for each $x \in \mathfrak{X}$ and if the linear bounded operator,

$$R(\lambda)x = \int_0^\infty \exp(-\lambda t)S(t)x\, dt, \tag{2.1}$$

defined for $\Re(\lambda) > w_0$ is such that $\lim_{\lambda\to\infty} \lambda R(\lambda)x = x$ for each $x \in \mathfrak{X}$. For semigroups of both classes $\lim_{t\to 0+} t^{-1}[S(t) - I]x \equiv Ax$ exits for a set $\mathfrak{D}(A)$ dense in $\mathfrak{X}$. A is called the infinitesimal generator of $[S(t)]$. A is closed for semigroups of class (C_0). For semigroups of class $(0, A)$, A has a smallest closed extension $\bar{A}$, called the complete infinitesimal generator (c.i.g). In either case the resolvent $R(\lambda; \bar{A})$ exists and $R(\lambda) = R(\lambda; \bar{A})$ for $\Re(\lambda) > \omega_0$. Finally, if $x \in \mathfrak{D}(\bar{A})$, then $\lim_{t\to 0+} S(t)x = x$ and

$$d[S(t)x]/dt = \bar{A}S(t)x = S(t)\bar{A}x \tag{2.2}$$

for $t > 0$. If $[S(t)]$ is of class (C_0), then (2.2) holds for all $t \geqq 0$ with $S(0) = I$, by definition.

3. The two formulations of the abstract Cauchy problem given below are suggested by properties of semigroups of class (C_0) and $(0, A)$, respectively. Both of these formulations differ somewhat from the Cauchy problem originally considered by E. Hille.

ACP_1 (ACP_2).—Given a linear operator U with domain and range in $\mathfrak{X}$ and given an element $y_0 \in \mathfrak{X}$, find a function $y(t) = y(t; y_0)$ such that

(i) *$y(t)$ is strongly absolutely continuous and continuously differentiable in each finite subinterval of $[0, \infty)$ [or $(0, \infty)$];*

(ii) *$y(t) \in \mathfrak{D}(U)$ and $U[y(t)] = y'(t)$ for each $t > 0$;*

(iii) *$\lim_{t\to 0+} y(t; y_0) = y_0$.*

The two problems differ only in the following respect: In ACP_1, $y'(t; y_0)$ is continuous for $t \geqq 0$, whereas in ACP_2, $y'(t; y_0)$ is continuous only for $t > 0$. In both cases $\int_0^1 \|y'(t; y_0)\|\, dt < \infty$.

We can now give a precise statement of our main results.

THEOREM 3.1. *Let U be a closed linear operator with dense domain and nonvacuous resolvent set. Suppose that for each $y_0 \in \mathfrak{D}(U)$ there is a unique solution to ACP_1. Then U generates a semigroup $[S(t)]$ of class (C_0) such that $S(t)y_0 = y(t; y_0)$ for all $y_0 \in \mathfrak{D}(U)$.*

THEOREM 3.2. *Let U be a closed linear operator with dense domain and such that $\|R(\lambda; U)\| = 0(1/\lambda)$ as $\lambda \to \infty$. Suppose that for each $y_0 \in \mathfrak{D}(U)$ there is a unique solution to ACP. Then U generates a semigroup $[S(t)]$ of class $(0, A)$ such that $S(t)y_0 = y(t; y_0)$ for all $y_0 \in \mathfrak{D}(U)$.*

Perhaps the most critical assumption in the above theorems is that the solution of the ACP be unique for each $y_0 \in \mathfrak{D}(U)$. That this assumption is not unduly restrictive follows from

THEOREM 3.3. *If $\bar{A}$ is the c.i.g. of a semigroup of class $(0, A)$, then ACP_2 has a unique solution for each $y_0 \in \mathfrak{D}(\bar{A})$.*

Proof: The existence of a solution to the ACP_2 for each $y_0 \in \mathfrak{D}(\bar{A})$ follows from the relation (2.2). If there were two solutions for a given $y_0 \in \mathfrak{D}(\bar{A})$, then the difference $z(t)$ would be a null solution of ACP_2 that is $\lim_{t\to 0+} z(t) = \theta$. It suffices to show that $z(t) = \theta$ for all $t > 0$. Now

$$\frac{d}{d\sigma}[S(t-\sigma)z(\sigma)] = S(t-\sigma)(-\bar{A} + \bar{A})z(\sigma) \equiv \theta, \qquad 0 < \sigma < t.$$

Consequently, $S(t-\beta)z(\beta) = S(t-\alpha)z(\alpha)$ for $0 < \alpha < \beta < t$. Passing to the limit first as $\alpha \to 0+$ and then as $t \to \beta+$, we obtain $z(\beta) = S(\beta)z(0) = \theta$ for all $\beta > 0$.

Theorems 3.1 and 3.2 are very closely related, and it is possible to establish certain lemmas which are applicable to the proofs of both.

LEMMA 3.1. *Let U be a closed linear operator and suppose for each $y_0 \in \mathfrak{D}(U)$ that there exists a unique solution to ACP_2. Then for each $t > 0$ there exists a constant C_t such that*

$$\|y(t; y_0)\| + \|U[y(t; y_0)]\| \leqq C_t(\|y_0\| + \|Uy_0\|). \tag{3.1}$$

Proof:[3] We first introduce a new norm in $\mathfrak{D}(U)$, namely, $\|x\|' = \|x\| + \|Ux\|$, and denote the so topologized vector space by $\mathfrak{D}(U)'$. The operator U being closed, it follows that $\mathfrak{D}(U)'$ forms a Banach space. We also introduce a second auxiliary space $\mathfrak{C}'$ consisting of all strongly continuous functions $x(t)$ defined on $(0, \infty)$ to $\mathfrak{D}(U)'$ such that $\int_0^1 \|x(\sigma)\|' \, d\sigma < \infty$. We define a metric in $\mathfrak{C}'$ by

$$d[x(\cdot), y(\cdot)] = \sum_{n=0}^{\infty} 2^{-n} \frac{\|x(\cdot) - y(\cdot)\|_n}{1 + \|x(\cdot) - y(\cdot)\|_n},$$

where

$$\|x(\cdot)\|_0 = \int_0^1 \|x(\sigma)\|' \, d\sigma,$$

$$\|x(\cdot)\|_n = \sup\,[\|x(t)\|'; n^{-1} < t < n], \qquad n \geqq 1.$$

It is easy to show that $\mathfrak{C}'$ is a complete metric space. We now consider the transformation T on $\mathfrak{D}(U)'$ to $\mathfrak{C}'$ which takes $y_0 \in \mathfrak{D}(U)'$ into the corresponding solution of the ACP_2, that is, $y(t; y_0)$. Because of the uniqueness assumption, it follows

that T is linear on $\mathfrak{D}(U)'$. We shall now show that T is closed. To this end suppose that $y_n \to y_0$ in $\mathfrak{D}(U)'$ and that $T(y_n) \to v(\cdot)$ in $\mathfrak{S}'$. In particular, $y'(t; y_n) = U(y(t; y_n)] \to U[v(t)]$ in the $\mathfrak{X}$ norm uniformly in each closed subinterval of $(0, \infty)$ and $\int_0^1 \|y'(\sigma; y_n) - U[v(\sigma)]\| \, d\sigma \to 0$. As a consequence, the relation $y(t; y_n) = y_n + \int_0^t y'(\sigma; y_n) \, d\sigma$ tends in the limit to $v(t) = y_0 + \int_0^t U(v(\sigma)] \, d\sigma$. It follows that $v(t)$ is a solution of the ACP_2 with initial value y_0, and the uniqueness now implies that $v(t) = y(t; y_0)$. This proves that T is closed. Hence by the closed-graph theorem[4] we see that T is continuous. Continuity at $y_0 = \theta$ implies the existence of constants C_n, $n \geqq 1$, such that $\|y(\cdot\,; y_0)\|_n \leqq C_n \|y_0\|'$ for $\|y_0\|'$ sufficiently small. However, since both sides of this inequality are homogeneous in y_0, the result is valid for all $y_0 \in \mathfrak{D}(U)'$; this establishes (3.1).

LEMMA 3.2. *Let U be a closed linear operator and suppose for each $y_0 \in \mathfrak{D}(U)$ that there exists a unique solution to ACP_2. If $y_0 \in \mathfrak{D}(U^2)$, then*

$$y(t; Uy_0) = U[y(t; y_0)] = y'(t; y_0), \qquad t \geqq 0. \tag{3.2}$$

Proof: If $y_0 \in \mathfrak{D}(U^2)$, then $y(t; Uy_0)$ is strongly absolutely continuous and $U[y(t; Uy_0)] = y'(t; Uy_0)$ is Bochner integrable (relative to the $\mathfrak{X}$ norm) in each interval $(0, \beta)$. Consequently,

$$y(t; Uy_0) = Uy_0 + \int_0^t y'(\sigma; Uy_0) \, d\sigma = Uy_0 + U[\int_0^t y(\sigma; Uy_0) \, d\sigma].$$

Setting $z(t) = y_0 + \int_0^t y(\sigma; Uy_0) \, d\sigma$, we see that $z(t)$ is continuously differentiable for $t \geqq 0$ and that $z'(t) = y(t; Uy_0) = U[z(t)]$. Thus $z(t)$ is a solution of the ACP_2 with initial value y_0, and the uniqueness implies that $z(t) = y(t; y_0)$. The relation (3.2) now follows.

We come now to the proof of Theorems 3.1 and 3.2. We define the operator $S(t)y_0 = y(t; y_0)$ on $\mathfrak{D}(U)'$ to itself for $t > 0$. The uniqueness of the solution to the ACP_2 implies that $S(t)$ is linear, and Lemma 3.1 implies that $S(t)$ is bounded. Conditions (i) and (ii) of ACP_2 imply that $S(t)$ is continuous in the strong-operator topology of $\mathfrak{D}(U)'$ for $t > 0$. Further, uniqueness implies that $y(t_1 + t_2; y_0) = y[t_1; y(t_2; y_0)]$ for each $y_0 \in \mathfrak{D}(U)$ and $t_1, t_2 > 0$. In terms of the operator $S(t)$, this means that $S(t_1 + t_2) = S(t_1)S(t_2)$, so that $S(t)$ defines a semigroup of linear bounded operators on $\mathfrak{D}(U)'$, strongly continuous for $t > 0$. Choose $\lambda_0 \in \rho(U)$. Then for any $x \in \mathfrak{D}(U)$ there exists a $y \in \mathfrak{D}(U^2)$ such that $x = \lambda_0 y - Uy$. By Lemma 3.2, $S(t)Uy = US(t)y$, and hence

$$S(t)x = \lambda_0 S(t)y - S(t)Uy = \lambda_0 S(t)y - US(t)y.$$

Setting $M_1 = \max.\,(1, |\lambda_0|)$, we have

$$\begin{aligned} \|S(t)x\| &\leqq M_1 (\|S(t)y\| + \|US(t)y\|) = M_1 \|S(t)y\|' \\ &\leqq M_1 \|S(t)\|'(\|y\| + \|Uy\|). \end{aligned}$$

Now $y = R(\lambda_0; U)x$, so that $\|y\| \leqq \|R(\lambda_0; U)\| \, \|x\|$ and $\|Uy\| = \|\lambda_0 y - x\| \leqq M_2 \|x\|$, where $M_2 = M_1 \|R(\lambda_0; U)\| + 1$. Combining this with the above relation, we finally obtain

$$\|S(t)x\| \leqq 2M_1 M_2 \|S(t)\|' \|x\| \tag{3.3}$$

for all $x \in \mathfrak{D}(U)$. Since $\mathfrak{D}(U)$ is assumed dense in $\mathfrak{X}$, the operator $S(t)$ can now be

uniquely extended to be linear and bounded on $\mathfrak{X}$ itself. The relation (3.3) remains valid for the extended $S(t)$, and it is easy to show that $S(t)$ becomes a semigroup of linear bounded operators on $\mathfrak{X}$, strongly continuous for $t > 0$. By definition, $S(t)y_0 = y(t; y_0)$ for each $y_0 \in \mathfrak{D}(U)$ and $t > 0$.

Proof of Theorem 3.1.—Condition (i) of ACP_1 implies that $\lim_{t\to 0+} y(t; y_0)$ and $\lim_{t\to 0+} U[y(t; y_0)]$ exist in the $\mathfrak{X}$ norm for each $y_0 \in \mathfrak{D}(U)$. Therefore, $\lim_{t\to 0+} S(t)y$ exists in $\mathfrak{D}(U)'$ for each $y_0 \in \mathfrak{D}(U)'$, and, by the uniform boundedness theorem, $\|S(t)\|'$ remains bounded near $t = 0$. According to (3.3), $S(t)$ is likewise bounded near $t = 0$. Condition (iii) of ACP_1 asserts that $\lim_{t\to 0+} \|S(t)x - x\| = 0$ for each $x \in \mathfrak{D}(U)$, and, applying the Banach-Steinhaus theorem, we see that the same is true for all $x \in \mathfrak{X}$. Consequently, $S(t)$ is of class (C_0). It remains to show that $A = U$. If $y_0 \in \mathfrak{D}(U)$, then by condition (i) of ACP_1 we see that $S(t)y_0 = y(t; y_0)$ is in particular differentiable at $t = 0$; thus $y_0 \in \mathfrak{D}(A)$. Also $y(t; y_0) \to y_0$ and $U[y(t; y_0)] = y'(t; y_0) \to Ay_0$, and this, together with the fact that U is closed, implies that $Uy_0 = Ay_0$. $Uy_0 = Ay_0$; in other words $U \subset A$. Consequently if $y_0 \in \mathfrak{D}(U)$ and $\Re(\lambda) > \omega_0$, then $\exp(-\lambda t)US(t)y_0 = \exp(-\lambda t)S(t)Ay_0$ is Bochner integrable on $[0, \infty)$ and hence $UR(\lambda; A)y_0 = \int_0^\infty \exp(-\lambda t)S(t)Ay_0 dt$ exists. Thus $UR(\lambda; A)y = AR(\lambda; A)y = \lambda R(\lambda; A)y - y$ for all $y \in \mathfrak{D}(U)$, and, since $\mathfrak{D}(U)$ is dense in $\mathfrak{X}$ and U is closed, we see that this relation holds for all $y \in \mathfrak{X}$. Since $R(\lambda; A)[\mathfrak{X}] = \mathfrak{D}(A)$, it follows that $\mathfrak{D}(U) \supset \mathfrak{D}(A)$ and hence that $A = U$. This concludes the proof of Theorem 3.1.

Proof of Theorem 3.2.—In the statement of Theorem 3.2 we have assumed that $\|R(\lambda; U)\| = 0(1/\lambda)$ as $\lambda \to \infty$. Thus, for $x \in \mathfrak{D}(U)$, $\|\lambda R(\lambda; U)x - x\| = \|R(\lambda; U)Ux\| \to 0$ as $\lambda \to \infty$, and, by the Banach-Steinhaus theorem, $\lim_{\lambda\to\infty} \lambda R(\lambda; U)x = x$ for all $x \in \mathfrak{X}$. Suppose that $y \in \mathfrak{D}(U)$. Then $y_n \equiv nR(n; U)y \in \mathfrak{D}(U^2)$ for n sufficiently large, $y_n \to y$, and $Uy_n = nR(n; U)Uy \to Uy$. According to Lemma 3.2, $US(t)y_n = S(t)Uy_n$; and, passing to the limit as $n \to \infty$, we see that $US(t)y = S(t)Uy$; here we have made use of the fact that U is closed. Finally, if $y_0 \in \mathfrak{D}(U)$, then $S(t)Uy_0 = US(t)y_0 = y'(t; y_0)$ is Bochner integrable near $t = 0$ by condition (i) of ACP_2. Now for $\lambda_0 \in \rho(U)$ and $x \in \mathfrak{X}$ there exists a $y \in \mathfrak{D}(U)$ such that $x = \lambda_0 y - Uy$. It follows that $S(t)x = \lambda_0 S(t)y - S(t)Uy$ is Bochner integrable near $t = 0$. We now define $R(\lambda)$ as in (2.1). It can be shown that $R(\lambda)$ defines a linear bounded operator for each λ, $\Re(\lambda) > \omega_0$. For $y_0 \in \mathfrak{D}(U)$ we have, as above,

$$R(\lambda)Uy_0 = \int_0^\infty \exp(-\lambda t)\, y'(t; y_0)\, dt,$$

and an integration by parts gives $R(\lambda)Uy_0 = -y_0 + \lambda R(\lambda)y_0$. Thus $R(\lambda)(\lambda I - U)y = y$ for all $y \in \mathfrak{D}(U)$. If λ is sufficiently large, then $\lambda \in \rho(U)$ and $(\lambda I - U)$ maps $\mathfrak{D}(U)$ onto $\mathfrak{X}$, from which it follows that $R(\lambda) = R(\lambda; U)$. Thus $\lambda R(\lambda)x = \lambda R(\lambda; U)x \to x$ as $\lambda \to \infty$ for each $x \in \mathfrak{X}$, and hence $[S(t)]$ is of class $(0, A)$. Finally, as we have already noted, $R(\lambda) = R(\lambda; \bar{A})$ for $\lambda > \omega_0$, where $\bar{A}$ is the c.i.g. This implies that $\bar{A} = U$, which concludes the proof.

* This paper was written under the sponsorship of the Office of Ordnance Research, U.S. Army, under contract DA-04-495-ORD-406.

† Present address, Department of Mathematics, Yale University, New Haven, Conn.

[1] E. Hille, "A Note on Cauchy's Problem," *Ann. Soc. polonaise math.*, **25**, 56–68 (1952); "Une généralisation du problème de Cauchy," *Ann. Inst. Fourier*, **4**, 31–48 (1953).

[2] See E. Hille, "Functional Analysis and Semi-groups," *Amer. Math. Soc. Coll. Lectures,*

(1948); and R. S. Phillips, "An Inversion Formula for Laplace Transforms and Semi-groups of Linear Operators," **59**, 325–356 (1954).

[3] The method of proof employed in Lemma 3.1 was suggested by some work of P. Lax on the continuous dependence of the solution of a linear partial differential equation on initial data.

[4] S. Banach, *Théorie des opérations linéaires*, Warszawa, 1932, p. 41.

NEW TYPES OF TRINOMIAL CONGRUENCE CRITERIA APPLYING TO FERMAT'S LAST THEOREM

BY H. S. VANDIVER

UNIVERSITY OF TEXAS, AUSTIN, TEXAS

Communicated Feburary 20, 1954

In another paper[1] the writer obtained a result which may be stated as follows:

THEOREM I. *If l is an odd prime and*

$$x^l + y^l + z^l = 0 \tag{1}$$

and $y \equiv 0 \pmod{l}$; $(xz, l) = 1$, the prime ideal $\mathfrak{p}$ in K (the algebraic field defined by $\zeta = e^{2i\pi/l}$) is a divisor of the ideal (p); p a rational prime;

$$c = \frac{N(\mathfrak{p}) - 1}{l},$$

where $N(\mathfrak{p})$ is the norm of $\mathfrak{p}$, further, $\mathfrak{p}$ belongs to an exponent[2] which is prime to l with $xyz \not\equiv 0 \pmod{\mathfrak{p}}$, then

$$\binom{c}{s} + \binom{c}{s+l} t^l + \binom{c}{s+2l} t^{2l} + \cdots \tag{2}$$

is divisible by p for $s = 1, 2, \ldots, l-1$, and $y/x = t$.

In another article[3] we developed these criteria further and obtained criteria independent of t which were expressed in determinant form. These were complicated, however, and seemed too difficult to handle. The problem then arose to obtain, if possible, the equivalent of some of the material used in the proof of Theorem I in a form which seemed simpler. Such a form was found by the writer in the year 1944 but is published here for the first time.

We have[1] (relation [5a]), with the conditions in Theorem I on $\mathfrak{p}$,

$$(x + \zeta^a y)^c \equiv 1 \pmod{\mathfrak{p}}. \tag{3}$$

Now let $p = 1 + cl$, that is, the ideal $\mathfrak{p}$ is of the first degree. Now the statement in (3) is equivalent to the statement that

$$x + \zeta^2 y \equiv \omega_a{}^l \pmod{\mathfrak{p}}, \tag{4}$$

$a = 1, 2, \ldots, p-1$, with ω_a some integer in K. We may select a rational integer g_p as a primitive root of the ideal $\mathfrak{p}$ such that

$$\mathfrak{p} = (g^c - \zeta, p), \tag{5}$$

where we set g for g_p during an argument where p is fixed (as in the work up to the

statement of Theorem II which follows). Then this gives, using (5),

$$g^{ac} \equiv \zeta^{a} \pmod{\mathfrak{p}},$$

and consequently, from (4),

$$x + g^{ac}y \equiv \omega_a{}^{l} \pmod{\mathfrak{p}}.$$

Further, since $\omega_a = b_0{}^{(a)} + \ldots + b_1{}^{(a)}\zeta + b_{l-2}{}^{(a)}\zeta^{l-2}$, with the b's rational integers,

$$\omega_a \equiv \sum_{i=0}^{l-2} b_i{}^{(a)}\zeta^{i} \equiv \sum_{i=0}^{l-2} b_i{}^{(a)}g^{ci} \pmod{\mathfrak{p}},$$

and substituting in (4) gives

$$x + g^{ac}y \equiv k_a{}^{l} \pmod{p} \tag{6}$$

for $a = 1, 2, \ldots, p-1$, where k_a is a rational integer and $(k_a, p) = 1$. From (1) we also have, using the given condition $(z, p) = 1$,

$$x + y = k_0{}^{l}; \qquad (k_0, p) = 1, \tag{6a}$$

where k_0 is a rational integer. Now g is also a primitive root of the rational prime p, and, since by definition $xyz \not\equiv 0 \pmod{p}$, we have integers r and s such that $g^{r} \equiv x$ and $g^{s} \equiv y \pmod{p}$ and $k_a \equiv g^{t_a} \pmod{p}$: (6) and (6$a$) then give, for $a = 0, 1, 2, \ldots, l-1$,

$$1 + g^{ac+s-r} \equiv g^{lt_a - r} \pmod{p}. \tag{7}$$

Now in another paper[4] the writer considered the congruence for a fixed i and j,

$$1 + g^{i+cx_1} \equiv g^{j+lx_2} \pmod{p}, \tag{8}$$

$0 \leq i \leq c-1; 0 \leq j \leq l-1$. We consider the problem of finding the distinct sets x_1 and x_2 in (7) with x_1 in the set $0, 1, \ldots, l-1$ and x_2 in the set $0, 1, \ldots, c-1$. We shall represent[5] the number of these distinct sets x_1, x_2 by $(i, j)_{cl}$, and use (i, j) when throughout an argument c and l are fixed. If e and f are any integers, we define $(e, f)_{cl}$ as $(|e|, |f|)_{cl}$, where $|e|$ represents the least positive or zero residue of e modulo c, and $|f|$ equals the best positive or zero residue of f modulo l. Evidently $(i, j) = (i + c\alpha, j + l\beta)$. We know (N, relation [4]), since l is odd, that

$$\sum_{j=0}^{l-1} (i, j) = l - \begin{cases} 0 \text{ if } i \neq c/2, \\ 1 \text{ if } i = c/2. \end{cases} \tag{9}$$

Now in the relation (7) of the present paper we have l congruences with values of $a = 0, 1, 2, \ldots, l-1$ with corresponding values t_a. In effect, this states that the number of solutions x_1, x_2 in (7) is, for the present particular case, equal to $(s - r, l - r)$ and that this number equals l.

We may then state:

THEOREM II. *If (1) is satisfied in non-zero integers, l is an odd prime with $y \equiv 0 \pmod{l}$ and $(xz, l) = 1$, the prime ideal $\mathfrak{p}$ in K belongs to an exponent prime to l, $\mathfrak{p}$ divides the ideal (p) where p is a rational prime of the form $1 + cl$, $xyz \not\equiv 0 \pmod{p}$, then there is an i and a j in (8) such that $(i, j)_{cl} = l$.*

In N we refer[3] to what we called the *two extreme cases* in connection with the (i, j), *first*, where a given i was such that (i, j) was unity for each j and the *second*,

when for a given i there exists a j such that $(i, j) = l$. The second case applies to our Theorem II. It follows from (9) also that when we state that $(i, j) = l$ for a particular i and j, then $(i, h) = 0$ for $h \not\equiv j \pmod{l}$.

In N, Mrs. Pearson and the writer examined the (i, j)'s for the values of $l = 3, 4, 5, 6, 7, 11$, and 13, and various primes p of the form $1 + cl$. It was found for $l > 5$ that for none of the 34 primes p tested did the second extreme case hold for any i and j. (The primes p were not selected as being necessarily such that a prime ideal divisor $\mathfrak{p}$ of (p) belongs to an exponent prime to l.) The largest p examined was 1123 with $l = 11$.

In N, we considered our trinomial congruence initially in the form

$$1 + av^c \equiv bw^m \pmod{p}, \tag{10}$$

where $abxy \not\equiv 0 \pmod{p}$, with p an odd prime. We signalized three problems involving (10), and we shall now consider the third one treated there; namely, "If in (10) a and b are fixed integers, $abxy \not\equiv 0 \pmod{p}$, m is fixed, and p increases together with c subject to the relation $mc = k(p - 1)$ with k a fixed integer, what can be said about the number of the sets of solutions x, y?" We shall now consider the relation of this to the problem of our present paper. If we take our relation (6) and divide through by x, noting that $(x, p) = 1$, the result may be written in the form

$$1 + g^{ac}\frac{y}{x} \equiv y^{l-1} h_a{}^l \pmod{p}, \tag{11}$$

where $h_a y \equiv k_a \pmod{p}$. We also set here $k = 1$ and $m = l$. It will be seen that (11) is of the same form as (10) except that y/x in (11) is not expressed explicitly in integral form. Owing to this fact, if we do express this y/x as congruent to an integer modulo p, then this integer will vary with each p. In view of the fact that what theory we have concerning these trinomial congruences nearly all involves the use of primitive roots as in (8) of this paper, we shall continue our investigations on this basis. In spite of the fact that a primitive root changes with each prime, it turns out that we shall be able to treat (11) when x and y are fixed and p varies over the various primes. This will be shown elsewhere.

In connection with our problem concerning (10) as applied to (1), we shall show that it is possible to select a class of primes p, for a certain (i, j), such that the congruence (10) is satisfied for $m = l$ (an odd prime) sets of values v^c, w^m. Set in (10), $lc = p - 1$, with $a = 1$ and $b = 1$ and select p so that it is the norm of a prime ideal $\mathfrak{p}$ with p a principal ideal (ω) such that

$$\omega = 1 + 2\lambda^i\theta,$$

$\lambda = \zeta - 1$, θ an integer in K. Also we can choose our primitive root so that (5) holds. With $a = 1$ and $b = 1$ in (10) and applying (5), we have, noting them that the possible incongruent values of v^c in (10) are congruent to $1, \zeta, \zeta^2, \ldots, \zeta^{l-1}$, modulo $\mathfrak{p}$,

$$\left(\frac{1 + \zeta^a}{\mathfrak{p}}\right) = 1 \tag{12}$$

for $a = 1, 2, \ldots, l-1$, where $(\beta/\mathfrak{p})$ is a power of ζ such that

$$\beta^{[N(\mathfrak{p})-1]/l} \equiv \left(\frac{\beta}{\mathfrak{p}}\right) \pmod{\mathfrak{p}}; \qquad [(\beta), \mathfrak{p}] = 1,$$

with β an integer in K. The relation (12) follows from the fact that $(1 + \zeta^a)$ is a unit in K and ω is primary.[6]

Again, since ω is primary, we have, by the law of reciprocity for lth powers,[6]

$$\left(\frac{2}{\mathfrak{p}}\right) = \left(\frac{2}{\omega}\right) = \left(\frac{1+2\lambda'\theta}{2}\right) = \left(\frac{1}{2}\right) = 1. \tag{12a}$$

This shows that, if (10) is written in the form (8) by employing primitive roots, after we set $m = l$, the resulting congruence (8) has l solutions x_1, x_2, that is,[7] $(0, 0) = l$. By a known theorem, the set of principal ideals $(1 + 2\lambda'\theta)$, if we let θ range over the integers in K, contains an infinity of prime ideals.[8]

In Theorem I we may replace x by z and z by x in its statement. Consequently, for p sufficiently large, there are two different quantities (i_1, j_1) and (i_2, j_2) which equal l. Since in the statement of Theorem I we have to replace y/x by y/z, and since from (11) we see that for p sufficiently large,

$$y/x \not\equiv y/z \pmod{p},$$

in fact, it is only necessary to take p as a prime $> y(x - z)$ to satisfy said condition, $x \not\equiv z$ in (1).

As noted earlier,[1] Theorem I refers only to the second case of Fermat's last theorem. In the other, or first, case, a criterion has been given[3] (p. 149, Theorem IV) based on methods analogous to those used in our original[3] proof of Theorem I. Using the methods of the present paper, we may give a companion theorem to our present Theorem II but which refers to the first case of the Fermat problem. However, the details will not be given here.

We have arrived at various consequences of the criteria mentioned here which we hope to be able to publish in further papers.

[1] "Summary of Results and Proofs on Fermat's Last Theorem (Third Paper)," these PROCEEDINGS, **15**, 45 (1929). Note that Theorem I refers to the *second case* of Fermat's Last Theorem as $y \equiv 0 \pmod{l}$, $zx \not\equiv 0 \pmod{l}$.

[2] We say that an ideal $\mathfrak{p}$ belongs to an exponent n when $\mathfrak{p}^n$ is a principal ideal but $\mathfrak{p}^d$ is not a principal ideal for $0 < d < n$.

[3] "Certain Congruence Criteria Connected with Fermat's Last Theorem," these PROCEEDINGS, **28**, 144–150 (1942).

[4] "On a New Problem Concerning Trinomial Congruences Involving Rational Integers," these PROCEEDINGS, **39**, 1278–1285 (1953). This paper will be referred to as "N."

[5] In his paper "On Cyclotomy and Extensions of Gaussian Type Quadratic Relations Involving Numbers of Solutions of Conditional Equations in Finite Fields," these PROCEEDINGS, **38**, 981–991 (1952), the writer used a notation (relation [1] of that paper) which for the form (8) of the present paper is equivalent to the statement that $(i, j)_{cl}$ is the number of solutions of the congruence

$$g^{i+cy_1} + g^{j+ly_2} \equiv 1 \pmod{p}$$

in the y's, but in N this number was designated by $[i, j]_{cl}$. *A reader who is referring to these several papers of the author should be careful to note these different notations.* Cf. note 6 of N.

[6] T. Takagi, "Ueber das Reciprocitätsgesetz in einem beliebigen algebräischen Zahlkörper," *J. Coll. Science, Tokyo Imp. Univ.*, **44**, 26 (1922), Satz 11.

[7] For $l = 3$, this is exemplified by the cases $p = 109$ and 127. For $l = 5$ an example is $p = 61051$. We have not determined the smallest prime of the latter type.

[8] E. Landau, "Ueber die Verteilung der Primideale in den Idealklassen eines algebräischen Zahlkörpers," *Math. Ann.*, **63**, 196–197 (1907); T. Takagi, "Ueber eine Theorie des relativ Abel'schen Zahlkörpers," *J. Coll. Science, Tokyo Imp. Univ.*, **41**, 19 (1920); H. Hasse, "Bericht ueber neuere Untersuchungen und Probleme aus der Theorie der algebräischen Zahlkörper," *Zahresbericht Deutsch. Math.-Vereinigung*, **35**, 1–55 (1926).

---

*DIRECTIONAL SENSITIVITY OF SINGLE OMMATIDIA IN THE COMPOUND EYE OF LIMULUS**

By Talbot H. Waterman

OSBORN ZOÖLOGICAL LABORATORY, YALE UNIVERSITY

Communicated by J. S. Nicholas, February 9, 1954

Compound eyes differ from those of other types, not because the mosaic quality of their image is imposed by a cellular retina; this is so in all eyes. Rather the compound eye is unique in having a dioptric apparatus which is itself also a mosaic, one made up of facets and lenses, each focusing light on different specific parts of a convex retina. Because of restricted angular apertures and differences in the orientation of the axes of these optic elements, single ommatidia do not individually survey the whole visual field of the eye. Instead, one such unit can be stimulated only by a fractional part of the total field, a small patch lying near its own optical axis. An image of this spot is the ommatidium's contribution to the over-all visual mosaic.

Accordingly, the directional sensitivity of an ommatidium in the compound eye will be one of its fundamental physiological properties. Despite such importance, however, practically no direct information is available on this point. The present report describes some measurements of this function initiated in the lateral eyes of *Limulus*. Their relationship to the mechanism of polarized-light sensitivity will be considered subsequently.[1]

Materials and Methods.—Lateral eyes of *Limulus* were prepared for study as described previously.[2] However, two modifications in the method of mounting the eye were required by the present experiments. (1) To avoid the problem of uncontrolled reflection and refraction when the eye was rotated, it was mounted in air and sealed with paraffin wax to a plastic holder in such a way that the optic nerve passed through a slot into the physiological saline, under which it was necessary to isolate single responding units. (2) The chamber was then placed on a circular stage mounted on a dissecting microscope. This permitted the eye's orientation in azimuth to be conveniently and accurately changed while the nerve and the recording electrodes could be observed closely. In preliminary experiments, an additional adjustment in elevation was included, but the necessity of using the saline bath in preparing fine bundles of axons prevented its effective use. Consequently, the

directional sensitivity of ommatidia in different anatomical planes had to be distinguished by mounting eyes in various positions on the holder.

The location and relative orientation of the ommatidia illuminated by the stimulating light were controlled by covering all but certain facets with opaque paint, as described elsewhere.[3] In this way axons from exposed ommatidia in the center of the eye, which were most favorable for these experiments, could be recognized at once. It will be seen in Figure 1 that in this region the optical axes of the corneal lenses are normal to the eye surface. In the more peripheral areas of the eye, as observed long ago by Exner,[4] marked deviations up to 30–40° from normality occur.

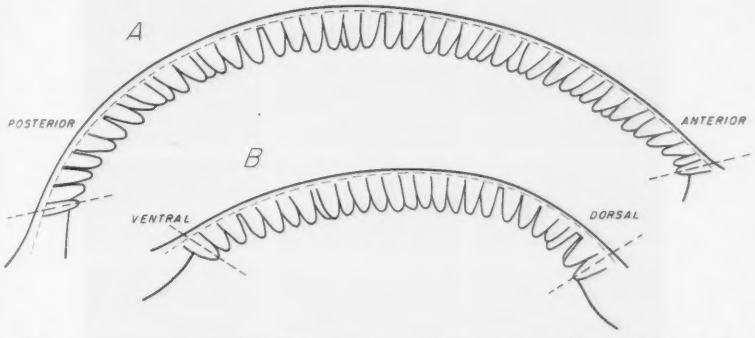

FIG. 1.—Camera lucida drawings of sections of the cornea and corneal lenses of a *Limulus* lateral eye (Juvenile ♀, 40 mm. prosoma length). The angle between the optical axes of marginal elements is 111° in the dorsoventral plane, 184° in the antero-posterior one. The actual visual field of the eye would be greater than this because of the large angular aperture of the ommatidia and the refractive effect of the cornea, where the optic axis is not normal to the surface. The angle of separation between the axes of neighboring facets varies from 4° to 15°. × 25.

The optical system used to stimulate the eye comprised the following. The source was a 6-volt, 18-ampere ribbon-filament microscope lamp, the image of whose uniform emitting surface was focused on a depolarizing filter comprising three layers of ordinary waxed paper mounted between optical blanks. In addition to the focusing lens, between the lamp and the depolarizer were a heat filter, a neutral Wratten filter for adjusting stimulus intensity in large steps, and two Polaroid sheets used for fine adjustments of transmitted intensity. A shutter was placed between the depolarizer and the *Limulus* eye preparation. Since this filter not only depolarized the light but also scattered it diffusely, it acted as the effective light source for stimulating the ommatidia. Consequently, the functional aperture for the stimulus was set by the diaphragm of the shutter. The geometry of the system was such that this evenly illuminated surface subtended an angle of 8° at the cornea. While further quantitative experiments need to be done with a smaller stimulating light source, it will be seen below that the directional sensitivity of the ommatidia in the *Limulus* eye is so broad that this factor did not greatly affect the present results.

The recording technique employed was standard, with tapered silver wire pickup electrodes, condenser-coupled amplifiers, and moving-film photography of the cathode-ray tube traces.

Experimental Results.—Three different ways of measuring directional sensitivity yielded the same qualitative conclusion, namely, that a single *Limulus* ommatidium is responsive over an astonishingly wide range of incident angles. In response to brief flashes of light, discharges could readily be obtained from a single unit in the optic nerve with the stimulus striking the cornea at angles as great as 30–40° away

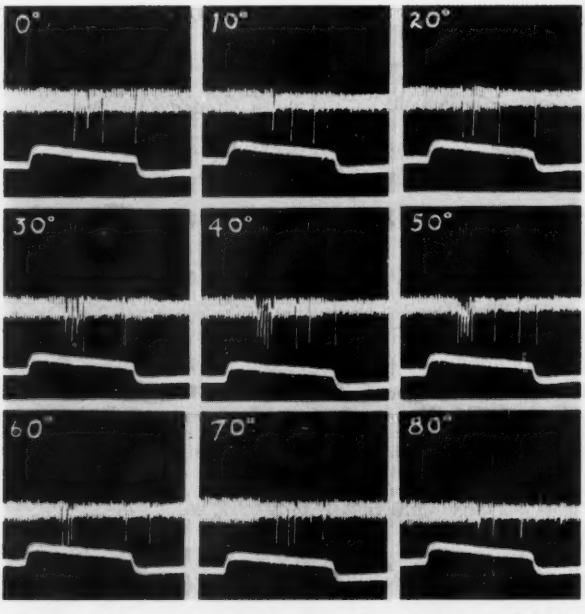

FIG. 2.—Responses of a single ommatidium to 0.7-second flashes of light striking the cornea at angles differing from an arbitrary 0° by the amounts indicated. The ommatidial axis, normal to the corneal surface, was located at about 45° on this scale. With light incident at −10° and at 90° or beyond, no response was obtained at the light intensity used throughout this series of exposures. The upper line of the records shows the single-unit optic-nerve impulses resulting from the flashes; the lower one indicates the occurrence and duration of the flash on the same temporal abscissa.

from the optical axis (Fig. 2). Within this range, of course, the number of impulses resulting from a flash of given intensity and duration decreased as the angle of incidence deviated to either side of the optical axis. Similarly, when a single ommatidium was stimulated with continuous illumination, a steady discharge was obtained over a wide range of incident angles (Fig. 3). Even in such light-adapted units, responses occurred with the light entering at 30° on either side of the optical axis.

The third way of measuring directional sensitivity, in addition to confirming the above results, also provided more quantitative data on the phenomenon. In this

case the threshold for a particular electrophysiological response was found as a function of the angle of stimulus incidence. The standard response was a series of four nerve impulses resulting from 1-second flashes given every 15 seconds. Stimulus intensity was adjusted at the various angles until this response was just obtained.

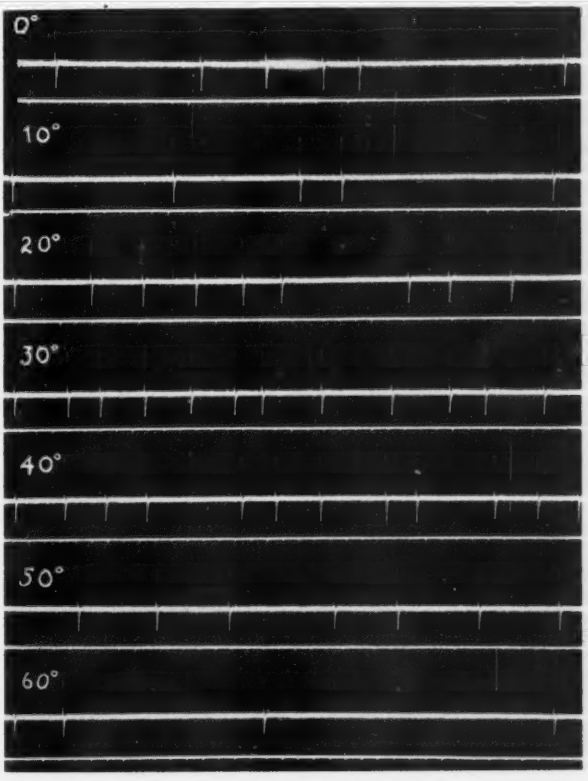

FIG. 3.—Responses of a single ommatidium to continuous illumination of constant intensity but of various incident angles, as indicated relative to an arbitrary 0°. The eye was light-adapted for 30 seconds at each of the different incident angles before the record was made. The optical axis of the unit, normal to the corneal surface, was between 30° and 35° on the arbitrary scale. The lower line in each strip indicates time in 0.1-second intervals.

Typical data secured in this way with either horizontal or vertical rotation of the eye are plotted in Figures 4 and 5. Obviously, the lowest thresholds were found close to the optical axes of the ommatidia, although high degrees of sensitivity extended out to 10–20° on either side. Responses frequently occurred with 40° deviations from the axis. Even larger angles were not unusual, and on one occasion the threshold in question at a 90° incident angle was less than 4 log units above the lowest threshold measured directly on the optic axis. Note that the effective apertures of a single ommatidium are twice the size of these deviations, carrying

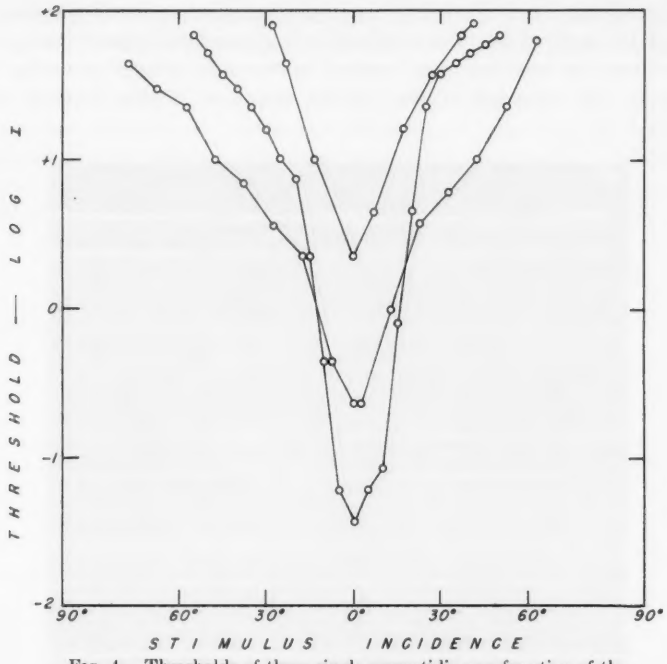

FIG. 4.—Thresholds of three single ommatidia as a function of the angle of stimulus incidence in the dorsoventral plane. The 0° represents the direction of the optical axis. The distinction between the absolute and angular sensitivities of the three units are the result of different orientations of their axes in the anteroposterior plane. In the most sensitive case the incident light made the smallest angle with the axis of the unit concerned. The threshold utilized to determine the light intensities plotted was a response consisting of 4 impulses in the optic axon involved. Zero on the light-intensity scale was 1 meter-candle.

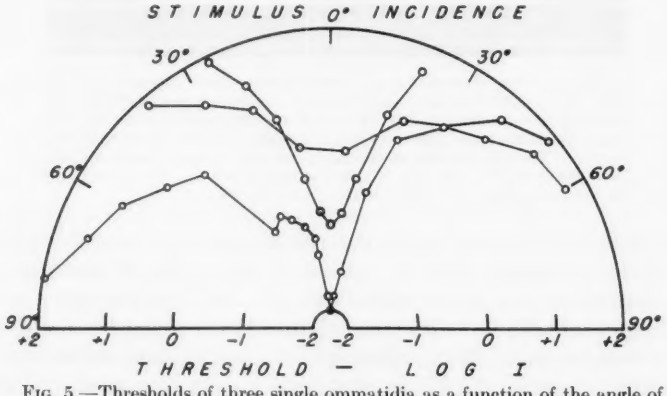

FIG. 5.—Thresholds of three single ommatidia as a function of the angle of stimulus incidence in the anteroposterior plane. Plotted on polar co-ordinates. Other features of the measurements and graph are the same as described for Figure 4.

high degrees of sensitivity to 40°, maximum to 180°! This should justify the statement made above that the 8° extent of the stimulus, although it could obscure fine detail, particularly close to the axis, had little influence on the general outcome of the measurements. Observe also that directional sensitivity appears to be similar in the two different anatomical planes.

Discussion.—The most important physiological consequence of the angular sensitivity of an ommatidium will be its relation to visual acuity. Nearly all the critical parameters of acuity, such as resolving power of the dioptric components,[5] the cellular mosaic of the retinulas concerned,[6] nervous integration,[7, 8] intensity and wave-length of illumination, contrast, and the spatial discrimination under test, were adequately controlled in the experiments. Consequently, the main factor to be discussed here is the relation of the observed directional sensitivity to the angle separating the ommatidial axes.

Measurements of corneal sections like those shown in Figure 1 indicate that the maximum angular separation of ommatidial axes is 15° or less, the minimum 4–5°. Clearly, then, the overlap in the visual fields of neighboring units must be extensive, since thresholds may not rise more than 2 log units with light incident 15° from the optical axis. Visual acuity might, for this reason, be less, and perhaps considerably so, than would be predicted from the number and angular separation of the ommatidia. Such a visual disability would fit in well with evidence from several other lines of work, suggesting that xiphosuran lateral eyes are degenerate structures.[3, 9] On the other hand, the high acuity of vertebrate retinas, despite marked overlap in the receptive fields of single optic-nerve fibers,[10] should provide a caveat against generalizing from optical or neurological data alone.

Experiments to determine the visual acuity of *Limulus* directly are now under way in the author's laboratory, with the optomotor response to moving stripes as a measure of visibility. Comparison of such behavioral results with the present electrophysiological data should further our understanding of the over-all problem.

This report may be briefly summarized by saying that measurements of directional sensitivity in single *Limulus* ommatidia show these units to respond over a broad range of incident stimulus angles. Although high sensitivity was centered within 10–20° of the optical axis, responses to reasonable light intensities were obtained 80–90° from this referent.

* These studies were aided by a contract between the Office of Naval Research, Department of the Navy, and Yale University, NR 163-091. A preliminary account of this work was presented at the December, 1953, meeting of the American Society of Zoölogists (*Anat. Rec.*, **117**, 566, 1953).

[1] T. H. Waterman, PROC. NATL. ACAD. SCI., **40**, 258–262 (1954).

[2] T. H. Waterman, *Science*, **111**, 252–254 (1950).

[3] T. H. Waterman and C. A. G. Wiersma, *J. Exper. Zoöl.* (in press).

[4] S. Exner, *Die Physiologie der facettirten Augen von Krebsen und Insecten*. Leipzig: Deuticke, 1891.

[5] H. B. Barlow, *J. Exper. Biol.*, **29**, 667–674 (1952).

[6] J. del Portillo, *Zeitschr. f. vergl. Physiol.*, **23**, 100–145 (1936).

[7] H. K. Hartline, H. G. Wagner and T. Tomita, *XIX Internat. Physiol. Cong., Abstr. of Communications*, pp. 441–442 (1953).

[8] B. Hassenstein, *Zeitschr. f. vergl. Physiol.*, **33**, 301–326 (1951).

[9] T. H. Waterman, *J. Morphol.* (in press).

[10] S. W. Kuffler, *J. Neurophysiol.*, **16**, 37–68 (1953)

*POLARIZED LIGHT AND ANGLE OF STIMULUS INCIDENCE IN THE COMPOUND EYE OF LIMULUS**

By Talbot H. Waterman

OSBORN ZOÖLOGICAL LABORATORY, YALE UNIVERSITY

Communicated by J. S. Nicholas, February 9, 1954

Evidence has been accumulating that the polarized-light sensitivities of various single ommatidia in the *Limulus* eye were different and that even the same ommatidium might show altered responses under changed conditions. Since the orientation of the optic axes of these units varies widely from one place to another in the eye, the most likely parameter concerned here seemed to be the incident angle of the stimulus. Consequently, experiments have been initiated to test the effect of this factor systematically. The general results so far obtained with nonpolarized

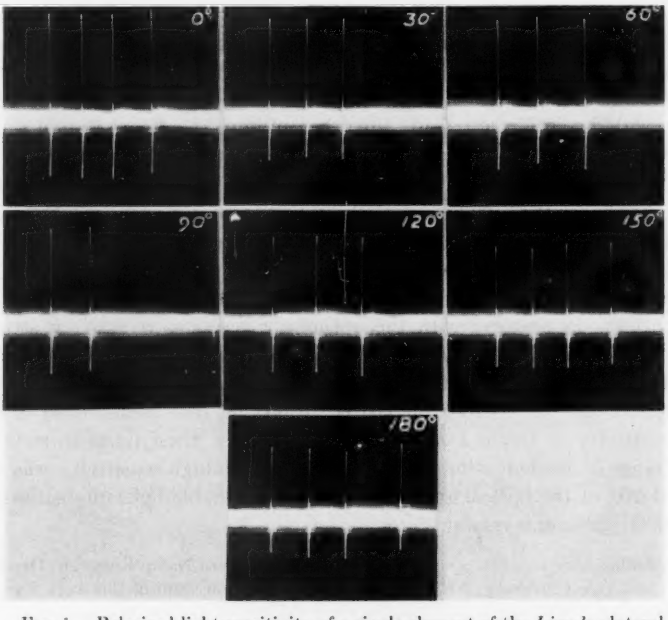

Fig. 1.—Polarized-light sensitivity of a single element of the *Limulus* lateral eye to 1-second flashes of plane-polarized light having a constant intensity but different planes of polarization, as indicated for each exposure. Maximum stimulation occurred near 0° and 180°; minimum near 90°, where only half as many optic-nerve impulses resulted from the standard flash.

light and a discussion of the significance of such directional sensitivity for visual acuity have already been presented.[1] This report describes the outcome of similar experiments with polarized light and considers the possible usefulness of such findings in determining the mechanism of polarized-light sensitivity.

Material and Methods.—The experimental preparation, the optical system, and the recording technique were all essentially the same as those used previously.[1] The only major innovation was the introduction of an adjustable Polaroid filter

between the depolarizer and the shutter of the optical system. This allowed systematic changes to be made in the plane of polarization of the stimulating light without affecting either its direction or its intensity. It is worth noting that the presence of the depolarizer may, by eliminating all stray polarization of whatever origin, be a highly important factor in the effective conduct of such experiments.

Results.—Whenever an electrophysiologically single element of the optic nerve showed marked sensitivity to the plane of polarization of the light illuminating the corresponding ommatidium, the following relationships were found. With con-

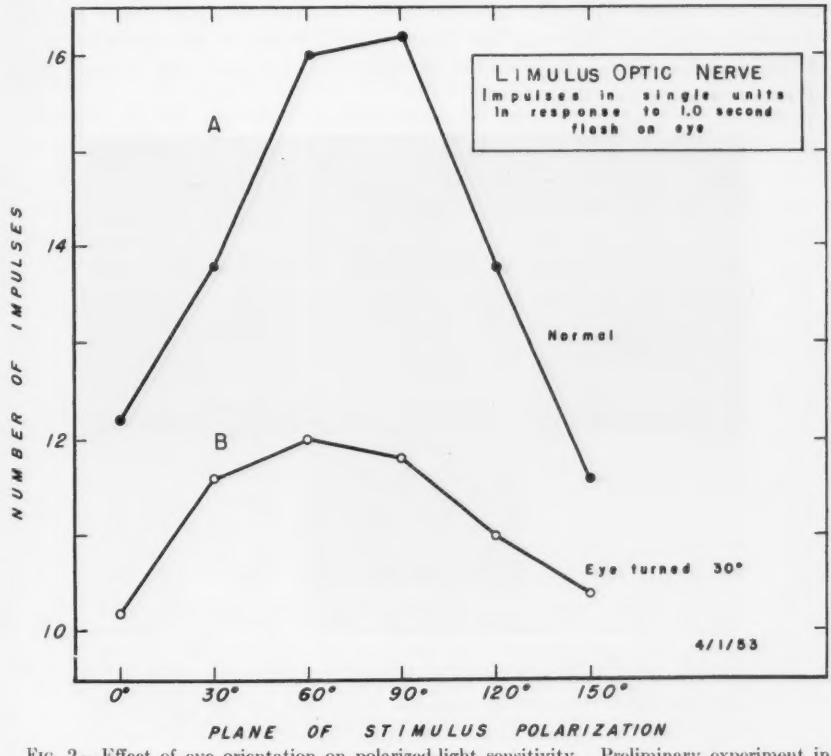

FIG. 2.—Effect of eye orientation on polarized-light sensitivity. Preliminary experiment in which it was observed that the high polarization sensitivity shown by a particular ommatidium in one orientation relative to stimulus incidence (curve *A*) was reduced to less than half when the eye was rotated (and hence the angle of incidence changed) through 30° (curve *B*).

tinuous illumination, the frequency of optic-nerve impulses varied systematically with the plane of polarization passing through maxima and minima 90° apart.[2] With flashes of light, provided that the unit was sufficiently sensitive to polarization, different numbers of action potentials resulted from exposures of the same intensity and duration but of different planes of polarization (Fig. 1). Again maxima and minima occurred 90° apart. Initially it was found that rotation of the eye into a new orientation relative to the stimulus could markedly affect the difference between these maximal and minimal points (Fig. 2).

Subsequently, the position of the optical axis of a single ommatidium was determined functionally by finding the most effective incident angle for the stimulus. This could be accomplished readily if only those ommatidia with optical axes normal to the corneal surface were used. Restriction of the excitable elements to the central region of the eye with opaque paint was an effective way of insuring this, as already described.[1] When the stimulus entered an ommatidium parallel to its optic axis and normal to the corneal surface, little or no polarized-light sensitivity was shown in the responses. This was true at least for flashes (Fig. 3); comparable measurements have not yet been made in continuous light because such long exposures are required to obtain adequate data. In contrast, units whose response to normally incident light was like that shown in Figure 3 gave evidence of an ability to analyze polarized light when the light stimulus had a large incident angle to the ommatidial axis (Fig. 4).

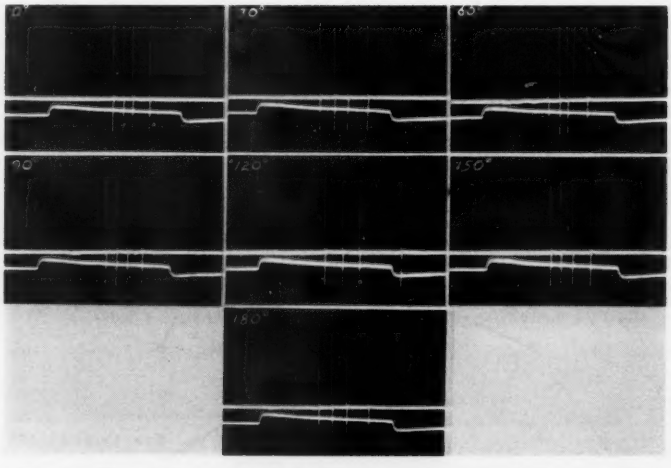

FIG. 3.—Virtual absence of polarized-light sensitivity of a single ommatidium exposed to 1.0-second flashes of plane-polarized light whose angle of incidence coincided with the optical axis of the unit. Successive flashes were of the same intensity, but differed in their planes of polarization as indicated. In all planes the same number of impulses were evoked by the flash. The upper line in each record shows the optic axon response to the flash, and the lower line the initiation and duration of the stimulus on the same temporal abscissa.

Discussion.—The discovery that the polarized-light sensitivity of the compound eye of at least one arthropod shows marked dependence on the angle of stimulus incidence may provide important clues to the mechanism of polarized-light perception.[3] In addition, the evidence that little or no polarization analysis was done on light incident at an angle close to the optical axis indicates the kind of further data necessary for the solution of this problem, namely, a precise quantitative knowledge of the relation between angle of incidence and polarization sensitivity. Before this additional information is obtained, however, it is possible to consider in how far the present findings illuminate the question of mechanism.

Perhaps the first possibility to be considered is that the polarization analyzer of arthropod eyes depends on simple reflection-refraction phenomena at the external

surface of the cornea. These would obviously be dependent on angle of incidence, as expressed quantitatively by Fresnel's formulas. Furthermore, the differential reflectance of light polarized in various planes is shown by these formulas to be negligible at low angles of incidence but to be considerable at the polarization angle and other large angles. As far as it goes, then, the present qualitative data might be accounted for on such a model. However, the possibility of complicating effects from internal reflection and refraction within the dioptric elements of the ommatidium should not be overlooked. Nor should one forget the angular sensitivity of the other possible means of analysis, which will be mentioned below.

Experimental data on the behavior of *Drosophila*,[4] *Daphnia*, water mites, and mosquito larvae[5] have, in fact, already been interpreted to indicate that refraction-reflection phenomena in the eyes form the sole sensory basis of their polarized-light behavior patterns. However, more extensive evidence is desirable for such a conclusion, particularly if it is to be considered of general applicability.

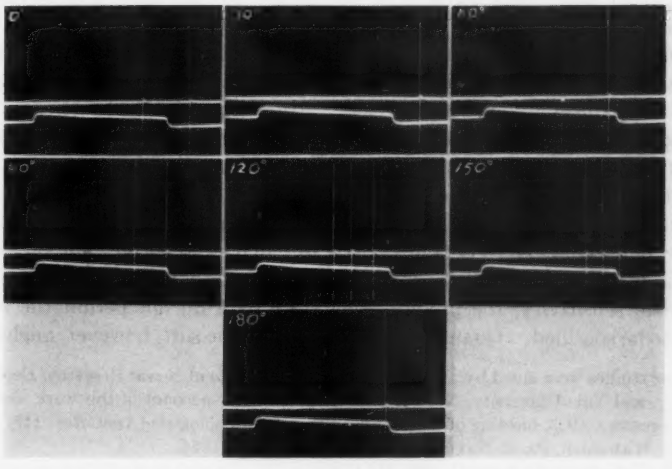

FIG. 4.—Polarized-light sensitivity of the same ommatidium whose responses are shown in Figure 3. In the present case the angle of stimulus incidence was 45° to the optical axis. As before, 1.0-second flashes of plane-polarized light were used as stimuli. These were of constant intensity but of different planes of polarization, as labeled. The lower line of the record again shows the initiation and duration of the stimulus on the same abscissa as the electrical activity of the optic axon responding.

Data from other arthropod eyes have been found to support different polarized-light analyzing mechanisms. Autrum and Stumpf[6] concluded from analysis of electroretinograms in the honey bee and in *Calliphora* that individual retinular cells act as detectors of the plane of light polarization. Since the stimulating light apparently entered the ommatidia responding at angles close to their optical axes, these data and their interpretation are incompatible with the kind of reflection-refraction mechanism proposed by Baylor and Smith and by Stephens, Fingerman, and Brown. Nor could Autrum and Stumpf's mechanism work in *Limulus*, since there each ommatidium ordinarily contributes but one active fiber to the optic nerve,[6] a restriction which apparently is not present in some insects.[7]

Optical indications that birefringence of the rhabdome in certain dipteran eyes is coupled with a mechanism for the extinction of one ray, somewhat as in a Nicol prism, were reported by Menzer and Stockhammer.[8] Finally, it has been suggested,[3, 7] partly on the basis of an analogy with the mechanism of seeing Haidinger's brushes by the human eye, that some dichroic element in the arthropod ommatidium might provide the required polarization analyzer.

Note that the three physical methods suggested as means of polarization analysis—(1) reflection-refraction, and birefringence with one ray eliminated either (2) by reflection or (3) by absorption—are all sensitive to the direction of the light rays traversing the system. Consequently, the present data do not alone provide adequate grounds for choice among them.

On the basis of the incomplete and conflicting information at present available on the mechanism of polarized-light analysis by arthropods, several points are clear. First, there may be various mechanisms operating in the eyes of different groups. Second, much more work needs to be done before any adequate general conclusions may be drawn. Finally, a quantitative investigation of the relation between polarization sensitivity and angle of light incidence is highly desirable in the case of *Limulus* at least. The precise shape of the function may indicate definitely the type of mechanism involved in this case. Also changes in the function induced by altering the refractive index of the external medium should show whether reflection-refraction at the corneal surface is a critical factor or not.

The main points of this report may be summarized as follows. The sensitivity of a *Limulus* ommatidium to polarized light varies with the angle of stimulus incidence. It is low or absent close to the optical axis and high at larger incident angles. Since the various suggested mechanisms for polarization analysis would all show some directional sensitivity, the present qualitative data do not permit the analytical device to be specified. Quantitative data of the same sort, however, might do so.

* These studies were aided by a contract between the Office of Naval Research, Department of the Navy, and Yale University, NR 163-091. A preliminary account of this work was presented at the December, 1953, meeting of the American Society of Zoölogists (*Anat. Rec.*, **117**, 566, 1953).

[1] T. H. Waterman, PROC. NATL. ACAD. SCI., **40**, 252–257 (1954).

[2] T. H. Waterman, *Science*, **111**, 252–254 (1950).

[3] T. H. Waterman, *Trans. New York Acad. Sci.*, **14**, 11–14 (1951).

[4] G. C. Stephens, M. Fingerman and F. A. Brown, Jr., *Ann. Entom. Soc. Amer.*, **46**, 75–83 (1953).

[5] E. R. Baylor and F. E. Smith, *Amer. Naturalist*, **87**, 97–101 (1953).

[6] T. H. Waterman and C. A. G. Wiersma, *J. Exper. Zoöl.* (in press).

[7] H. de Vries, A. Spoor and R. Jielof, *Physica*, **19**, 419–432 (1953).

[8] G. Menzer and K. Stockhammer, *Naturwiss.*, **38**, 190–191 (1951).

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general in accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph.*, Boston, **22**, 731–815 (1911); or, Wheeler, W. M., *Konigsberg, Schr. physik. Ges.*, **25**, 1–142 (1914).

MANUSCRIPTS will be accepted only from members of the Academy or from chairmen of the divisions of the National Research Council who will assume responsibility for the propriety and scientific standards of the paper and should be addressed to *Proceedings of the* NATIONAL ACADEMY OF SCIENCES, Attention Miss Mary D. Alexander, University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois. Articles of which a member of the Academy is author or co-author may extend to 10 printed pages, those by non-members to 5 printed pages; space above these limits will be billed to the author at $21.00 per page. Accompanying the manuscript should be a statement by the author or his representative that he assumes responsibility for the charges if incurred.

GALLEY PROOF will be sent, with the understanding that charges for author's corrections shall be billed to him. Author should therefore make final revisions on the typewritten manuscripts. Page proofs will not be sent.

REPRINTS should be ordered when proof is returned, on the reprint order form which will be inclosed with the proof. The order blank gives a tabular estimation of the cost of reprints; a copy will be supplied at any time upon request.

CONTENTS

Page

BIOCHEMISTRY.—THE INHIBITION OF HYDROGENASE BY NITRIC OXIDE . . . . . *By Alvin I. Krasna and D. Rittenberg* 225

BIOCHEMISTRY.—A BIOCHEMICAL AND PHARMACOLOGICAL SUGGESTION ABOUT CERTAIN MENTAL DISORDERS . . . . . *By D. W. Woolley and E. Shaw* 228

MATHEMATICS.—DYNAMIC PROGRAMMING AND A NEW FORMALISM IN THE CALCULUS OF VARIATIONS . . . . . *By Richard Bellman* 231

MATHEMATICS.—A CHARACTERIZATION OF TAME CURVES IN THREE-SPACE . . . . . *By O. G. Harrold, Jr., H. C. Griffith, and E. E. Posey* 235

MATHEMATICS.—CONFORMAL GEOMETRY AND ELEMENTARY PARTICLES . *By R. Ingraham* 237

MATHEMATICS.—APPROXIMATION BY BOUNDARY VALUES OF ANALYTIC FUNCTIONS . . . . . *By Jacqueline Penez* 240

MATHEMATICS.—A NOTE ON THE ABSTRACT CAUCHY PROBLEM . . . . *By R. S. Phillips* 244

MATHEMATICS.—NEW TYPES OF TRINOMIAL CONGRUENCE CRITERIA APPLYING TO FERMAT'S LAST THEOREM . . . . . *By H. S. Vandiver* 248

PHYSIOLOGY.—DIRECTIONAL SENSITIVITY OF SINGLE OMMATIDIA IN THE COMPOUND EYE OF *Limulus* . . . . . *By Talbot H. Waterman* 252

PHYSIOLOGY.—POLARIZED LIGHT AND ANGLE OF STIMULUS INCIDENCE IN THE COMPOUND EYE OF *Limulus* . . . . . *By Talbot H. Waterman* 258

age
225
228
231
235
237
240
244
248
252
258